Journal
of
The Franklin Institute
Devoted to Science and the Mechanic Arts

Vol. 232 NOVEMBER, 1941 No. 5

BENJAMIN FRANKLIN AS A SCIENTIST.*

BY

ROBERT A. MILLIKAN, Ph.D., Sc.D., LL.D.,
Director, Norman Bridge Laboratory of Physics, California Institute of Technology.

Benjamin Franklin is perhaps the only American in that relatively small group of men of any time or country who, without having been either the head of a state or a military hero, have yet gained so conspicuous a place in history that their names and sayings are known the world over. Although he lived two hundred years ago in what was then a remote corner of the earth, far from any of the centers of world influence, yet his name and traits are still so widely known that the following incident could happen. Of what other American names, save possibly those of Washington and Lincoln, could anything like this be said? Franklin would have found it as interesting and amusing as will this audience.

One evening about two and a half months ago I was being shown by a devout Brahmin through the great Hindu temple at Madura in Southern India. There were many hundreds of Indians wandering about through the scores of rooms in the huge structure. Occasionally one of them would fall prostrate on his face before the image of the elephant god or the monkey god, or one or another of the multitudinous forms assumed by Krishna or Siva or their

* Delivered in the Hall of The Franklin Institute, April 18, 1940. This was one of the papers in the "Meet Dr. Franklin" Conference.

(Note—The Franklin Institute is not responsible for the statements and opinions advanced by contributors in the JOURNAL.)

wives. In one of the rooms an expounder of the Yogi philosophy was passing around a printed page of directions for "the Yogi way of life" to a score or more of pupils. As I was looking on in the back of the room he approached and put one of the sheets into my hand. I glanced through it and saw that it was in English and consisted merely in a set of rules for physical exercise (mystically interpreted, however), for eating and sleeping and early rising, etc., etc. As I glanced down the page I read the injunction, "You are to spend some time each day, morning and evening, in reflection, as Benjamin Franklin did."

But while that incident reflects Franklin's world-wide fame it is quite clear that it has nothing to do with Benjamin Franklin as a scientist. May I introduce the consideration of that side of Franklin's character by quoting a paragraph from a short biography of Michelson which I published in the *Scientific Monthly* for January, 1939:

"It will probably be generally agreed that the three American physicists whose work has been most epoch-making and whose names are most certain to be frequently heard wherever and whenever in future years the story of physics is told are Benjamin Franklin, Josiah Willard Gibbs, and Albert A. Michelson. And yet the three have almost no characteristics in common. Franklin lives as a physicist because, dilettante though he is some times called, mere qualitative interpreter though he actually was, yet it was he who with altogether amazing insight laid the real foundations on which the whole superstructure of electrical theory and interpretation has been erected. Gibbs lives because, profound scholar, matchless analyst that he was, he did for statistical mechanics and for thermodynamics what Laplace did for celestial mechanics and Maxwell did for electrodynamics, namely, made his field a well-nigh finished theoretical structure. Michelson, pure experimentalist, designer of instruments, refiner of techniques, lives because in the field of optics he drove the refinement of measurement to its limits and by so doing showed a skeptical world what far-reaching consequences can follow from that sort of a process and what new vistas of knowledge can be opened up by it. It was a lesson the world had to learn. The results

of learning it are reflected today in the extraordinary recent discoveries in the field of electronics, of radioactivity, of vitamins, of hormones, of nuclear structure, etc. All these fields owe a large debt to Michelson, the pioneer in the art of measurement of extraordinarily minute quantities and effects."

In that paragraph I have tried to appraise Franklin's place among American scientists. Let me now express my personal judgment as to his place in world science. If I were asked to list by centuries the fourteen most *influential* scientists who have lived since Copernicus was born in 1473, I should bring forward the following names by centuries:

Century.	Name.	Nationality.
15th Century	Copernicus (1473–1543)	Pole
16th Century	Galileo (1564–1642)	Italian
17th Century	Newton (1642–1727)	English
	Huyghens (1629–1695)	Dutch
18th Century	Franklin (1706–1790)	American
	La Place (1749–1827)	French
19th Century	Faraday (1791–1867)	English
	Maxwell (1831–1879)	English
	Darwin (1809–1882)	English
	Fresnel (1788–1827)	French
	Pasteur (1822–1895)	French
	Gauss (1777–1855)	German
	Helmholtz (1821–1894)	German
	Volta (1748–1827)	Italian
	Willard Gibbs (1839–1903)	American

There will doubtless be those, especially among Europeans, who will say, "Why do you give Franklin so high a place when there were but seven of his eighty-four years, namely, from 1746 to 1753, in which he pursued science at all, also when he wrote, so far as I can discover, not a single scientific paper designed for publication in a scientific journal?" His private letters to his friend Peter Collinson which he never expected to be published at all, are practically the sole source of our knowledge of his scientific work. Even his own estimate of his scientific achievement was so small that in his autobiography he makes but casual reference to it.

The answer to the foregoing inquiry is that I have been guided in the placing of Franklin and all of the others on the

above list primarily by the significance of their contributions as measured by the influence they exerted in the development of our modern world. I have not been concerned at all with the erudition of the candidates, the profundity or extent of their scholarship, nor even by the magnitude and difficulty of the problems which they solved.

No one, however, can read these letters to Peter Collinson —I have had access to the fifth edition published in 1774— without being amazed by the fact that Franklin without any previous training whatever in either the technique or the history of physics and with almost no contact with what others were doing or had done, within two years of the time of his first experiment had acquired a keener insight into the fundamental nature of electrical phenomena, not merely than any one had acquired up to his time, but even than any of his successors acquired for the next hundred and fifty years, when, about 1900, the scientific world returned essentially to Franklin's views.

To justify this statement and to bring to light the extraordinary quality both of Franklin's physical insight and of his power of induction I shall make most of the remainder of this lecture consist of a few direct quotations from the Peter Collinson letters which the editor informs us were being printed "without waiting for the ingenious author's permission to do so."

The first letter, dated March 28, 1747, reads:

"To Peter Collinson, Esq; F. R. S. London

Philadelphia, March 28, 1747

"Sir,

"Your kind present of an electric tube,[1] with directions for using it, has put several of us on making electrical experiments, in which we have observed some particular phænomena that we look upon to be new. I shall therefore communicate them to you in my next, though possibly they may not be new to you, as among the numbers daily employed in those experiments on your side the water, 'tis probable some one or other has hit on the same observations. For my own part, I never was before engaged in any study that so totally engrossed my attention and my time as this has lately done;

for what with making experiments when I can be alone, and repeating them to my Friends and Acquaintance, who, from the novelty of the thing, come continually in crouds to see them, I have, during some months past, had little leisure for any thing else.

"I am, etc.

"B. Franklin."

Now as to some of the experiments themselves. The very first one of them, done within a few months of the time he first heard of electricity, contains the key to his invention of the lightning rod. Note from the following how skillfully and strikingly he arranges his electrostatic experiments by making the length of the suspension of the cork ball very long. After two hundred years of the development of electrostatics these experiments cannot be made more tellingly today than by setting them up and performing them *exactly* as Franklin directed nearly two hundred years ago.

He writes, "The first is the wonderful effect of pointed bodies, both in *drawing off* and *throwing off* the electrical fire. For example,

"Place an iron shot of three or four inches diameter on the mouth of a clean dry glass bottle. By a fine silken thread from the cieling, right over the mouth of the bottle, suspend a small cork-ball, about the bigness of a marble; the thread of such a length, as that the cork-ball may rest against the side of the shot. Electrify the shot, and the ball will be repelled to the distance of four or five inches, more or less, according to the quantity of Electricity. . . . When in this state, if you present to the shot the point of a long, slender, sharp bodkin, at six or eight inches distance, the repellency is instantly destroyed, and the cork flies to the shot. A blunt body must be brought within an inch, and draw a spark, to produce the same effect. To prove that the electrical fire is *drawn off* by the point, if you take the blade of the bodkin out of the wooden handle, and fix it in a stick of sealing-wax, and then present it at the distance aforesaid, or if you bring it very near, no such effect follows; but sliding one finger

[1] A straight three-foot glass tube as big as your wrist.

along the wax till you touch the blade, and the ball flies to the shot immediately."

Here is where he learned that his lightning rod had to have a good ground in order to work at all. He continues:

"To show that points will *throw off* as well as *draw off* the electrical fire, lay a long sharp needle upon the shot, and you cannot electrise the shot so as to make it repel the cork-ball. . . . Or fix a needle to the end of a suspended gun-barrel, or iron-rod, so as to point beyond it like a little bayonet; and while it remains there, the gun-barrel, or rod, cannot by applying the tube to the other end be electrised so as to give a spark, the fire continually running out silently at the point."

I can find no evidence that prior to Franklin the electrical properties of points had been discovered at all. He continues:

"The repellency between the cork-ball and the shot is likewise destroyed, (1) by sifting fine sand on it; this does it gradually; (2) by breathing on it; (3) by making a smoke about it from burning wood; (4) by candle-light, even though the candle is at a foot distance: these do it suddenly. . . . The light of a bright coal from a wood fire; and the light of a red-hot iron do it likewise; but not at so great a distance.

"The light of the sun thrown strongly on both cork and shot by a looking-glass for a long time together, does not impair the repellency in the least. This difference between fire-light and sun-light is another thing that seems new and extraordinary to us." [2]

The insight shown in the three lines, of the foot note below in which he correctly makes particle carriers (ions, we now call them) from the match do the discharging while sunlight produces no ions and therefore does not discharge, is unbelievably penetrating for a date two hundred years back, though the conception of neutral particles being first attracted and then repelled is of course definitely wrong.

The next experiment, with its interpretation, is probably the most fundamental thing ever done in the field of electricity. Get it exactly in Franklin's words:

[2] "This different Effect probably did not arise from any difference in the light, but rather from the particles separated from the candle, being first attracted and then repelled, carrying off the electric matter with them."

"1. A person standing on wax, and rubbing the tube, and another person on wax drawing the fire, they will both of them (provided they do not stand so as to touch one another) appear to be electrised, to a person standing on the floor; that is, he will receive a spark on approaching each of them with his knuckle.

"2. But if the persons on wax touch one another during the exciting of the tube, neither of them will appear to be electrised.

"3. If they touch one another after exciting the tube, and drawing the fire as aforesaid, there will be a stronger spark between them than was between either of them and the person on the floor.

"4. After such strong spark, neither of them discover any electricity.

"These appearances we attempt to account for thus: We suppose, as aforesaid, that electrical fire is a common element (we now call 'electrical fire' electrons), of which every one of the three persons abovementioned has his equal share, before any operation is begun with the tube. *A*, who stands on wax and rubs the tube, collects the electrical fire from himself into the glass; and his communication with the common stock being cut off by the wax, his body is not again immediately supply'd. *B* (who stands on wax likewise), passing his knuckle along near the tube, receives the fire which was collected by the glass from *A*; and his communication with the common stock being likewise cut off, he retains the additional quantity received. . . . To *C*, standing on the floor, both appear to be electrised: for he, having only the middle quantity of electrical fire, receives a spark upon approaching *B*, who has an over quantity; but gives one to *A*, who has an under quantity. If *A* and *B* approach to touch each other, the spark is stronger, because the difference between them is greater: After such touch there is no spark between either of them and *C*, because the electrical fire in all is reduced to the original equality. If they touch while electrising, the equality is never destroy'd, the fire only circulating. Hence have arisen some new terms among us: we say *B* (and bodies like circumstanced) is electrised *positively; A*, *negatively*. Or rather, *B* is electrised *plus; A*, *minus*.

And we daily in our experiments electrise bodies *plus* or *minus*, as we think proper.—To electrise *plus* or *minus*, no more needs to be known than this, that the parts of the tube or sphere that are rubbed, do, in the instant of the friction, attract the electrical fire, and therefore take it from the thing rubbing: the same parts immediately, as the friction upon them ceases, are disposed to give the fire they have received, to any body that has less."

The next two long letters are taken up largely with what he calls "M. Muschenbroek's wonderful bottle," accidentally discovered in Leyden one year earlier, 1746, now known as the Leyden jar, and with explaining all such effects just as we do today in terms of the opposite charges or the inner and outer coats.

Thus, to use his exact words, "At the same time that the wire and top (inside coat) of the bottle is electrified positively or plus the bottom (outside coat) of the bottle is electrified negatively or minus, in exact proportion: i.e., whatever quantity of electrical fire is thrown in at the top an equal quantity goes out at the bottom." And "Again, when the bottle is electrised, but little of the electrical fire can be drawn out from the top by touching the wire unless an equal quantity can at the same time *get in* at the bottom. Thus, place an electrised bottle on clean glass or dry wax and you will not, by touching the wire get out the fire from the top."

These chapters, too, contain the uncannily clever experiment of showing, just as we do today, that the charge resides in or on the dielectric. How many of us realize that the familiar class-room experiment of removing the coats of a Leyden jar and touching each of them, then putting them back again, and after that getting a strong spark by connecting the replaced coatings with a wire was devised by Benjamin Franklin in 1749.

Again, he says, "There is one experiment more which surprises us, and is not hitherto satisfactorily accounted for; it is this: Place an iron shot on a glass stand, and let a ball of damp cork, suspended by a silk thread, hang in contact with the shot. Take a bottle in each hand, one that is electrified through the hook, the other through the coating: Apply the

giving wire to the shot, which will electrify it *positively*, and the cork shall be repelled; then apply the requiring wire, which will take out the spark given by the other; when the cork will return to the shot: Apply the same again, and take out another spark, so will the shot be electrified *negatively*, and the cork in that shall be repelled equally as before. Then apply the giving wire to the shot, and give the spark it wanted, so will the cork return: Give it another, which will be an addition to its natural quantity, so will the cork be repelled again: And so may the experiment be repeated as long as there is any charge in the bottles. *Which shews that bodies having less than the common quantity of electricity, repel each other, as well as those that have more.*"

In that last sentence Franklin states clearly that matter which had lost its normal amount of electricity was self repellent. In modern terms the atom is neutral when it has its full complement of electrons. When any of these are removed the nuclei repel one another.

The next paragraph is added merely to illustrate the amusing and dramatic side of Franklin's character.

"Chagrined a little that we have been hitherto able to produce nothing in this way of use to mankind; and the hot weather coming on, when electrical experiments are not so agreeable, it is proposed to put an end to them for this season, somewhat humorously, in a party of pleasure, on the banks of Skuylkil. Spirits, at the same time, are to be fired by a spark sent from side to side through the river, without any other conductor than the water; an experiment which we some time since performed, to the amazement of many. A turkey is to be killed for our dinner by the *electrical shock*, and roasted by the *electrical jack*, before a fire kindled by the *electrical bottle:* when the healths of all the famous electricians in England, Holland, France, and Germany are to be drank in *electrified bumpers*, under the discharge of guns from the *electrical battery.*"

In some of these letters, notably the fifth, Franklin goes off into long and incorrect speculations as to the difference between the terms "electric bodies per se" and "non electric bodies." But this adds to, rather than subtracts from my own appreciation of him, for no human being could possibly

have seen correctly all the elements of a huge and thus far completely unexplored field, and his wrong steps give him opportunity to show his greatness by the way he goes to work to discover and to admit his error. Thus, he writes as follows:

"*Query*, Wherein consists the difference between an *electric* and a *non-electric* body?

"*Answer*. The terms electric *per se*, and non-electric, were first used to distinguish bodies, on a mistaken supposition that those called electrics *per se*, alone contained electric matter in their substance, which was capable of being excited by friction, and of being produced or drawn from them, and communicated to those called non-electrics, supposed to be destitute of it: For the glass, etc., being rubb'd, discover'd signs of having it, by snapping to the finger, attracting, repelling, etc. and could communicate those signs to metals and water. . . . Afterwards it was found, that rubbing of glass would not produce the electric matter, unless a communication was preserved between the rubber and the floor; and subsequent experiments proved that the electric matter was really drawn from those bodies that at first were thought to have none in them. Then it was doubted whether glass and other bodies called *electrics per se*, had really any electric matter in them, since they apparently afforded none but what they first extracted from those which had been called non-electrics. But some of my experiments shew that glass contains it in great quantity, and I now suspect it to be pretty equally diffused in all the matter of this terraqueous globe. If so, the terms *electric per se*, and *non-electric*, should be laid aside as improper: and (the only difference being this, that some bodies will conduct electric matter, and others will not) the terms *conductor* and *non-conductor* may supply their place."

Without doubt the most profound paragraphs in all of Franklin's letters are the following, written in 1749:

"1. The electrical matter consists of particles extremely subtile, since it can permeate common matter, even the densist metals, with such ease and freedom as not to receive any perceptible resistance.

"2. If any one should doubt whether the electrical matter passes through the substance of bodies, or only over and along their surfaces, a shock from an electrified large glass jar, taken through his own body, will probably convince him.

"3. Electrical matter differs from common matter in this, that the parts of the latter mutually attract, those of the former mutually repel each other. Hence the appearing divergency in a stream of electrified effluvia.

"4. But though the particles of electrical matter do repel each other, they are strongly attracted by all other matter.

"5. From these three things, the extreme subtility of the electrical matter, the mutual repulsion of its parts, and the strong attraction between them and other matter, arise this effect, that, when a quantity of electrical matter is applied to a mass of common matter, of any bigness or length, within our observation (which hath not already got its quantity) it is immediately and equally diffused through the whole.

"6. Thus common matter is a kind of spunge to the electrical fluid. And as a spunge would receive no water if the parts of water were not smaller than the pores of the spunge; and even then but slowly, if there were not a mutual attraction between those parts and the parts of the spunge; and would still imbibe it faster, if the mutual attraction among the parts of the water did not impede, some force being required to separate them; and fastest, if, instead of attraction, there were a mutual repulsion among those parts, which would act in conjunction with the attraction of the spunge. So is the case between the electrical and common matter.

"7. But in common matter there is (generally) as much of the electrical as it will contain within its substance. If more is added, it lies without upon the surface, and forms what we call an electrical atmosphere; and then the body is said to be electrified."

In these paragraphs Franklin states with great succinctness what later became known as the Franklin one-fluid theory, and after 1900 was known as the electron theory. In his day, and for 150 years thereafter, it received very scant consideration in the old world, and the so-called two-fluid theory of Aepinus, put forward a little later, was universally taught in

text books the world over up to the triumph of the electron theory in 1897 under the active leadership of J. J. Thomson, who himself pointed out that this electron theory was in essential particulars a return to the theory put forth by Franklin in 1749. For Franklin's electrical matter consisted of extremely subtle mobile *particles* (now called negative electrons), which in order to make matter exhibit its common or neutral properties had to be present in each kind of matter (we now say in each kind of atom; but the atomic theory had not been formulated in 1749) in a particular number, an increase in which number made it exhibit electrification of one sign, a decrease an electrification of the opposite sign. In Franklin's theory only one kind of electrical matter was mobile; the other sign of electrification appeared when the mobile kind was removed so that it could no longer neutralize the effect of the opposite kind *which inhered in the immobile part of matter* (i.e. in the nucleus).

The Franklin theory was mathematically identical with the two-fluid theory, but while the former was a definite and profound *physical* theory the latter was a hold-over from mediæval mysticism. It came from the age of the so-called "imponderables"—an imponderable or weightless heat theory, the caloric—and the imponderable electric fluids. Such vague, tenuous, contradictory ideas were ill at home in the highly realistic, practical mind of Franklin. They were justified, like Faraday's lines of magnetic force, as analytical conveniences but not as physical realities. Franklin introduced a definite physical theory which rendered unnecessary such fantastic conceptions as two weightless and hence non-existent fluids introduced for purely *ad hoc* purposes, and then told to destroy each other, also for *ad hoc* purposes.

Let us now return to Franklin's discussion of points and their properties of throwing off or drawing off the electrical fire. He says, very modestly and wisely:

"These explanations of the power and operation of points, when they first occurred to me, and while they first floated in my mind, appeared perfectly satisfactory; but now I have written them, and considered them more closely, I must own I have some doubts about them; yet, as I have at present nothing better to offer in their stead, I do not cross them out:

for even a bad solution read, and its faults discovered, has often given rise to a good one, in the mind of an ingenious reader."

Then in the next paragraph note how clearly he sees the necessity of eliminating unnecessary hypotheses, i.e., he adopts the scientific principle of "minimum hypothesis."

"Nor is it of much importance to us, to know the manner in which nature executes her laws; it is enough if we know the laws themselves. It is of real use to know that china left in the air unsupported will fall and break; but *how* it comes to fall, and *why* it breaks, are matters of speculation. It is a pleasure indeed to know them, but we can preserve our china without it."

He then describes some discharging effects of points conducted on a larger scale than he had before attempted, and in a later paper dated November 7, 1749, he enumerates all the known points of resemblance between lightning and electricity, and concludes with the comment:

"The electric fluid is attracted by points. We do not know whether this property be in lightning, but since they agree in all points in which we can compare them, it is not improbable that they agree likewise in this. Let the experiment be made."

In June, 1752 he made it, carrying out in a shed with his son the experiment which he describes as follows in his letter of October 19, 1752, to Peter Collinson.

"As frequent mention is made in public papers from Europe of the success of the Philadelphia experiment for drawing the electric fire from clouds by means of pointed rods of iron erected on high buildings, etc. it may be agreeable to the curious to be informed that the same experiment has succeeded in Philadelphia, though made in a different and more easy manner, which is as follows:

"Make a small cross of two light strips of cedar, the arms so long as to reach to the four corners of a large thin silk handkerchief when extended; tie the corners of the handkerchief to the extremities of the cross, so you have the body of a kite; which being properly accommodated with a tail, loop, and string, will rise in the air, like those made of paper; but

this being of silk is better to bear the wet and wind of a thunder gust without tearing. To the top of the upright stick of the cross is to be fixed a very sharp pointed wire, rising a foot more above the wood. To the end of the twine, next the hand, is to be tied a silk ribbon, and where the silk and twine join, a key may be fastened. This kite is to be raised when a thunder-gust appears to be coming on, and the person who holds the string must stand within a door or window, or under some cover, so that the silk ribbon may not be wet; and care must be taken that the twine does not touch the frame of the door or window. As soon as any of the thunder clouds come over the kite, the pointed wire will draw the electric fire from them, and the kite, with all the twine, will be electrified, and the loose filaments of the twine will stand out every way, and be attracted by an approaching finger. And when the rain has wet the kite and twine, so that it can conduct the electric fire freely, you will find it stream out plentifully from the key on the approach of your knuckle. At this key the phial may be charged; and from electric fire thus obtained, spirits may be kindled, and all the other electric experiments be performed, which are usually done by the help of a rubbed glass globe or tube, and thereby the sameness of the electric matter with that of lightning completely demonstrated."

In a further letter written in September, 1753, he says: "In September 1752 I erected an iron rod to draw the lightning down into my house, in order to make some experiments on it." He carried on these experiments for some months to learn whether the clouds were positively or negatively electrified, and after many trials he says,

"I concluded that the clouds are *always* electrified *negatively*, or have always in them less than their natural quantity of the electric fluid.

"Yet notwithstanding so many experiments, it seems I concluded too soon; for at last, June the 6th, in a gust which continued from five o'clock P.M. to seven, I met with one cloud that was electrified positively, though several that passed over my rod before, during the same gust, were in the negative state."

The foregoing shows what most commendable scientific care he took in his experiments and what caution he used in drawing conclusions.

But he did not stop with making scientific experiments. His active and practical mind was not satisfied until he had applied it to the useful end of the invention of the lightning rod, as indicated in the first paragraph of the letter of October 19, 1752, quoted above.

After his definite proof of the identity of lightning and electricity he was recognized by the most distinguished English scientists by being elected to the Royal Society, and was presented for the year 1753 the Copley medal of the Society, the highest honor within the gift of the world's most illustrious scientific body.

After its presentation at the Franklin Institute I sent the article to Mr. I. Bernard Cohen of the Harvard Library, and received from him the following reply, which I am very glad to have presented along with the article because it enables the reader to gain a little better view of the breadth of Franklin's interests, and also to avoid some minor misunderstandings.

HARVARD LIBRARY
CAMBRIDGE, MASSACHUSETTS
November 30, 1940

PROFESSOR ROBERT A. MILLIKAN

Dear Sir:

Thank you very much for your kindness in allowing me to read your essay on Benjamin Franklin. I think it very interesting and I am very glad that it is to be published.

There is nothing wrong in any of your statements, although I think that two of them are misleading. First, you write of Franklin's work in electricity and only that. It is true, of course, that Franklin did his most significant work in electricity, but his scientific range was considerably broader. Thus, he did pioneer work in locating the position of the Gulf Stream, he discovered that northeast storms come from the southwest, i.e., they travel in the direction opposite to that in

which they blow. He also made the first measurements of heat absorption with regard to colour, and was the first to investigate the conductivity of different substances with regard to heat. Similarly, he did other things in medicine (ventilation, the cause of colds, the invention of a flexible metal catheter) and advanced the cause of hospitals. And no small part of his contribution to science took the form of the advancement of it,—by forming the American Philosophical Society, by providing the means for publication of scientific works, by encouraging younger scientists (such as Priestley, etc.), and by spreading scientific information. This means only that in your essay you should, in my opinion, mention that you are concerned with Franklin and Electricity, and not write of Franklin and electricity as if it were Franklin and science.

Second, I cannot agree with your statement that Franklin wrote no article to be published in a scientific journal. We are misled by the form of his writing. An examination of the *Philosophical Transactions* of the eighteenth century shows that most of the articles were letters in form, addressed by a non-member of the society to a member, or by a member to the President. We know now that Franklin's letters were read at the meetings of the Royal Society and were discussed in various printed articles and pamphlets. This is one of my own finds, one not sufficiently brought out by any other writer. Thus the letters which Franklin wrote to Collinson were of two kinds: one was personal and never read to the Royal Society members, the other was public and was so read. This type of letter was the eighteenth-century equivalent of the "note" which one sends to a journal today. In this light, your statement is somewhat misleading. Further, there are some articles such as the "Loose Thoughts on a Universal Fluid" read at the American Philosophical Society and published in their Proceedings.

For the rest I found your article clear and refreshing. It is pleasant for me to find a first rank scientist like yourself stating firmly and definitely that Franklin was a scientist. If he did not give many years to science, that was not his fault; and we know that as he had to give more and more of his time to politics he lamented the fact more and more. He always

considered himself a scientist doing necessary work in politics, and at eighty-one he could ask his friend Ingenhousz to rejoice with him that he was a "freeman" after fifty years of public service and that now he could go back to America and resume his experiments, if the English had not destroyed his equipment.

Besides, as you say so clearly, we judge a man by his achievement. What does it matter whether Franklin spent six years or twenty years at his work? Newton is not great because he put so many work-hours into his discoveries, but because his discoveries were of great importance. Thus, no one censures Newton because he gave most of his time to alchemy, theology, and affairs of the mind.

Yours sincerely,

I. BERNARD COHEN

Iceless Ice. (*Compressed Air Magazine*, Vol. 46, No. 7.)—Most of us have heard of dry ice and have probably carried home a carton of ice cream packed in this solidified carbon dioxide to keep it frozen. But hot ice, so called because it reaches a temperature of as much as 78° F. is another matter. It is not a refrigerant, but otherwise has the characteristics of real ice and is being used in place of it in indoor skating rinks. Mark ice, as it is named after its inventor, John Markman, is a mineral product that is applied in liquid form. It sets almost immediately after spraying, it is claimed, making it popular with hotels where ice-skating revues are a feature of the entertainment. With the program finished, the floor can be quickly resurfaced, if necessary, for dancing. There is a rink of this kind in the Persian Room of the Sir Francis Drake Hotel in San Francisco, Calif. It measures 19 × 30 feet and has a 3/4 inch coat weighing 3½ tons, or approximately 2½ tons more than genuine ice of the same dimensions and thickness would weigh.

R. H. O.

More and Better Plywood.—The plywood industry is on the threshold of a new production era, according to a statement to a Senate committee by Paul H. Appleby, Undersecretary of Agriculture. "Output," he said, "increased from 200 million square feet in 1932 to 700 million square feet in 1936 owing mainly to the inherently fine qualities of plywood as a structural material—high strength for its weight, for example. But now plywood far superior to that formerly used is available by virtue of improved plastic adhesives that are impervious to water. Furthermore, the properties of wood and plywood alike can now be modified almost at will by impregnation with resin-forming materials followed by heat and pressure. This new type of plywood, waterproof and capable of being compressed and molded to form odd shapes and integral units, lends itself to almost unlimited development. In aircraft, for example, it can be used in the construction of wings, leading edges of wings, wing tips, fuselages, ailerons, flooring, bomb bay doors, instrument panels, etc. Its use for such purposes would tend to relive the aluminum alloy situation and at the same time would bring into the defense program an industry whose resources of labor and materials have not yet been tapped to the extent possible."

R. H. O.

THE DIRECT MEASUREMENT OF LIGHTNING CURRENT.

BY

JOHN WILSON FLOWERS, Ph.D.,

High Voltage Engineering Laboratory, General Electric Company, Pittsfield, Massachusetts.

ABSTRACT.

Two cathode-ray oscillographs have been designed and employed to record the current in the natural lightning stroke. Three years of operation in the tower of the world's tallest structure, the Empire State Building in New York City, have given many records of strokes to this structure. Automatic and coördinated operation together with sufficient magnitude and time range of recording have enabled current-time characteristics of the direct stroke to be obtained. One instrument gives the low-magnitude, long-time current record of each stroke, while the other yields the high-magnitude, short-time components. Coverage from 50 to 200,000 amperes and from one microsecond to one second is provided with the records related properly in order and time.

Many physical measurements have been made with the purpose of obtaining a better understanding of lightning and its effects. However, most measurements can be classified as indirect and the remaining few as inherently incomplete from the standpoint of fundamental lightning characteristics. Perhaps the most fundamental is that of the current in the stroke channel at the point of contact with the ground or other terminal object, and, needless to say, the most elusive, as the indirect and incomplete methods indicate. Viewed as a whole, much important information about the actual current has been revealed, yet much that is desirable has been lacking, and some erroneous ideas have been created. With a unique station available for the study of lightning current, this accumulated information has been a guide for the design of current recording equipment.

The Empire State Building in New York City is struck by lightning many times during the season when thunderstorms occur.[1] Because of the extreme height of this building as compared to surrounding structures, the tower top, which is over 1250 feet above the ground level, is a focal point for lightning. The ideal condition exists for the location of

measuring apparatus close to the contact point of the lightning stroke, and the complete current furnished by the stroke is available for study in contrast to various conditions which have existed in other experiments.

Where measurements are made remote from the contact point of the stroke, many difficulties enter when attempting to interpret physical characteristics. For a traveling wave with current passing over a relatively short path, distortion caused by various conditions is likely to alter the wave so that it becomes difficult of interpretation. This situation exists with transmission line studies where reflections, flash-over, unknown impedances, corona and other factors make the deduction of the original current wave a virtual impossibility. Field strength measurements in the vicinity of lightning activity are limited more than seems to have been indicated. Unless the true rate of change of field is completely known throughout the discharge, calculations of current magnitudes and polarity effects are unreliable, if not impossible. Charge transported cannot be determined without a knowledge of the complete field distribution over the region beneath the cloud both before and after the discharge.

The Empire State Building Investigations begun in 1935 have continued each year, and equipment developed and used through the 1937 season has already been described.[2,3] Nineteen long-time oscillographic records were obtained in 1937 by means of a galvanometer oscillograph.[1,3] No cathode-ray oscillograms were recorded mainly because of obscurities of the phenomena to be recorded and difficulties of laboratory testing which simulate actual lightning conditions. In 1938 a successful stage was reached and cathode-ray oscillograms were obtained of the current in the direct lightning stroke. Two newly developed cathode-ray oscillographs were employed. The methods of automatic control, the reliability of operation, and the coördination of the records obtained have proved the fundamental correctness of the method of measurement. Storms in 1938, 1939, and 1940 have given much information on the electrical characteristics of lightning. Both current magnitudes and wave shapes of many direct strokes have been measured.

DESIGN CONSIDERATIONS.

The cathode-ray oscillograph appeared to offer the most complete electrical information possible. By its use many valuable measurements have been made of surges produced by lightning on transmission lines,[4,5] and conductors exposed to the electric[6] and magnetic fields [7,8] of lightning discharges. References cited are by no means complete, but serve here as examples.

The choice of cathode-ray equipment depends upon many factors of which some of the more important are recording or writing speed, method of operation, and nature and usefulness of the records obtained. That high writing speed is needed has long been accepted, but little consideration by past investigations has been given to the method of recording other than the hope that upon one film there might be one oscillogram of value. No attempt has been made to separate and determine the time between important events or the order of occurrence. In view of the complex nature of the lightning stroke, it is absolutely necessary to provide a means of separating and establishing an order of surge records. This complexity has been indicated in many ways. Norinder [6] has recorded many rapid surges occurring throughout a lightning stroke. Photography with moving film or lenses has shown the duration of light from a stroke to be in some instances of the order of a second and a half with recurrent light flashes of varying characteristics.[1,9,10] It has also been shown that the recurring surges in the same lightning stroke can cause repeated flashovers of high voltage transmission lines with special protective devices.[11] Thus it was evident that recording apparatus needed to respond to extremely short time events yet be capable of covering each event which might occur over a relatively long time. It was also fairly evident, and to be expected, that in general high currents would be associated with short time. At least two oscillographs were thus necessary to cover the time and magnitude range of currents.

A high-speed oscillograph capable of recording all of the fast, high-current surges in a stroke was one necessary instrument. The universal choice for the detailed study of lightning transients has been some form of the continuously-

pumped oscillograph with internal photographic film for the reason that the oscillographs available were of this type. Various modifications have been made for giving improved performance such as the beam relay [5,6,12] and grid-controlled hot-cathode type [13] and many other improvements. While all instruments used are unsurpassed for high-speed recording, this general type has serious limitations in operation and recording. The problem of a high-speed instrument capable of recording many subsequent surges has been realized. The nearest approach to a solution seems to have been in the use of an oscillatory sweep [14] used by many investigators. While by this means repeated surges are recorded, they are superimposed upon each other without any order. It should be noted that when surges are measured on live transmission lines, the first surge can cause a power fault lasting until cleared and following heavy discharges in the same stroke may produce relatively reduced potential effects. This fact may have prevented much mixing of records which have been obtained on transmission lines where the method has been widely used. This problem of an oscillograph which can record at the necessary high speed each individual high current surge as a separate oscillogram of known time occurrence has been successfully solved by the high-vacuum-sealed type cathode-ray tube [15] together with moving film photographic recording.

Since a high-speed instrument is not suitable for providing a continuous record of low-magnitude slowly varying currents, another oscillograph must be designed accordingly. This low-speed oscillograph must give one oscillogram for each entire stroke and cover in time the complete current history. Where the detail is unresolved because of short times or high magnitudes, then the high-speed oscillograph must provide the detail. The requirements of this low-speed instrument are also fulfilled by a design of the general type of the high-speed oscillograph with suitable modifications.

Much has been written concerning lightning current magnitudes, particularly the higher values. No method used can be termed "precise" because of unknown factors which enter into each measurement. Laboratory experiments with high currents have permitted estimates by comparison of

heating and destructive effects.[16] Lichtenberg figure methods[17] have been widely used in the early stage of surge studies but are of limited value here. Because of simplicity, the magnetic link[18] has yielded information from a wide application.[19,20] Recent further developments have given a magnetic recorder capable of time resolution to a limited extent.[21] The general opinion can be taken from many sources that currents as high as 200,000 amperes can be expected and this value has been used as the upper limit for this investigation. Much less is probably known about the lower limit of lightning currents. However, various imposed conditions in the electrical arrangement have limited measurements by cathode-ray tubes to a value above 50 amperes. Photographic methods have shown luminosity existing below this current measuring limit,[1] and recent experiments in this laboratory indicate that, except for some conditions existing in high-speed photography, light causing photographic recording indicates simultaneous current flow.

The source of voltage to which the oscillographs are connected determines much about the general design. For the measurement of high current, this source must be very low in impedance and as low in voltage as is practical. Currents of the order of 100,000 amperes impose special requirements and restrictions upon the current-carrying parts of the measuring circuit. The current shunt, which is the source of voltage for the oscillographs, must be designed for this high current. Hence, if physical size is to be limited and certain electrical characteristics maintained, the voltage which appears across this shunt must be held within some limit. The maximum useful voltage from the standpoint of measurement is that required to give full scale deflection of the cathode-ray beam and is about 1000 volts for a high-voltage cathode-ray tube. Lower values cannot be used because amplification is not practical, and, while high voltage could be used for initiating purposes, such voltages would be very much higher than 1000 volts if ordinary methods of control such as three-electrode gap arrangements were used. Such voltages do not appear practical from the standpoint of any shunt or voltage-dropping device from which they can be obtained. Other methods of obtaining a high voltage, such

as might be achieved by a probe or gap, would involve high-current gaps in the main current circuit.

The hot-cathode type cathode-ray tube can be properly controlled by potentials lower than the minimum values which the tube will measure. Thus, potentials from the shunt to be measured are sufficient for control and the oscillographs can be adjusted in sensitivity to record all magnitudes which yield measurements. Of much importance is the fact that the low-magnitude voltages which produce initiation are the same voltages which are to be measured. It then becomes possible to select by proper electric circuits only surge voltages of importance to a high-speed oscillograph in both magnitude and time characteristics. The importance of this possibility is more apparent in consideration of the wide variations in the current-time characteristics of the lightning stroke. Without the ability to select suitable portions of the stroke for high-speed recording, the records would be hopelessly mixed, to say nothing of the quantities of records which would be unimportant for analysis by a high-speed oscillograph.

The recording speed of the sealed-off, hot-cathode type tube appears to be sufficient for any phenomenon so far found in the direct lightning stroke. For each oscillograph, coördinated control circuits and photographic recording methods have been designed. Similar drums carrying identical amounts of photographic film and rotating at the same speeds have enabled a comparison of records of the same disturbance in the proper time relations. However, because of the difference of operation, records obtained are different in appearance as will be evident.

THE HIGH-CURRENT CIRCUIT.

The heavy current of the lightning stroke is conducted through the building by way of the metal framework and the network of wires and pipes found in the building and leading to it. The lightning-current circuit of interest here comprises the lightning rod, a heavy copper cable, the high-current shunt, and the steel frame of the building. Before 1939 a vertical radiator of the National Broadcasting Company served as the lightning rod with the shunt inserted at a nodal

point as described.[2] When new radiators for the transmission of television programs were installed, a separate lightning rod designed specifically for the measurement of lightning current was incorporated. Limitations in the shunt design and

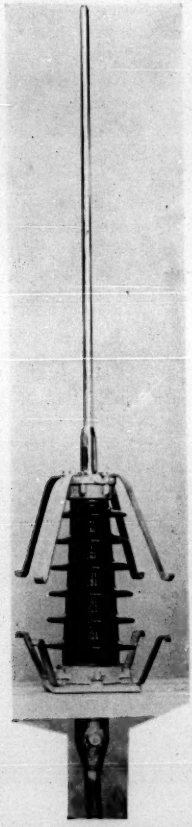


FIG. 1. Empire State Building lightning rod and cable terminal assembly.

difficulties with flashovers near the top of the antenna were eliminated by this change.

The top of the lightning rod is 36 feet above the tower roof and intercepts practically all strokes to the building. The rod which is four feet in length is supported by a high-

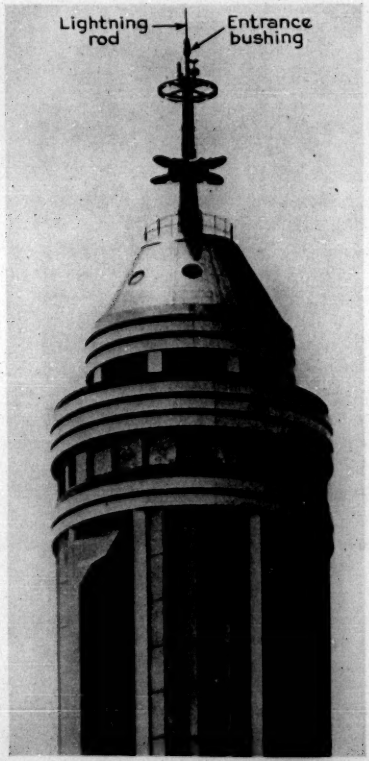


FIG. 2. View of the Empire State Building Tower showing rod and bushing mounted above high-frequency radiators.

voltage entrance bushing mounted in the end of a cylindrical structure which extends 30 feet above the roof. An oil-filled high-voltage cable 34 feet in length is attached to the bushing and passes down into the space below the roof. The bushing assembly before installation is shown in Fig. 1 and as it appears in a view of the tower top in Fig. 2. High-voltage insulation is necessary at the bushing and top portion of the cable even though a maximum of about 1000 volts appears at the shunt, because the surge impedance of the cable becomes a factor for currents with rapid rates of rise. Impulse insulation of about 200 Kv. is provided, and it should be pointed out that high-voltage conditions will always exist whether it is attempted to convey the current to the point of measurement or measure at a convenient distance from the current path. In the latter case the measuring equipment is subjected to high-voltage relative to its surroundings and constitutes the worst possible condition for oscillograph measurements as will be indicated later. The insulation provided by the bushing and cable has been found sufficient to convey the current the necessary 34 feet.

The cable leads to the high-current shunt 4 feet below the roof. Lightning current after passing through the shunt enters the metal framework of the building. Structural symmetry provides a radial distribution to the supporting members of the tower and the metal walls. This distribution much reduces the value of disturbing magnetic fields from which the cathode-ray tubes must be shielded.

THE HIGH-SPEED OSCILLOGRAPH.

Current through the shunt produces a potential which performs two functions. The potential applied to the control circuits initiates the cathode-ray beam and the sweep. Secondly, it causes a deflection of the beam proportional to voltage magnitude. By means of a high-speed lens and a sensitive photographic film, a single trace upon the fluorescent screen can be permanently recorded. With the film in motion, with certain circuit requirements fulfilled, and with a short persistence screen used, each trace upon the screen is recorded as a separate oscillogram. By virtue of the motion, oscillograms are separated and related in order and time. During

a stroke, initiation and recording may occur as many times as there are suitable current surges. The first initiation of the control circuits trips a relay circuit which gives an indication of operation upon a recording clock and automatically begins shifting the film to a new exposure in readiness for a subsequent stroke.

For transient measurements by means of the cathode-ray oscillograph, the external connections involving the apparatus are extremely important, as the manner of connecting influences the distribution of surge current in all wires and

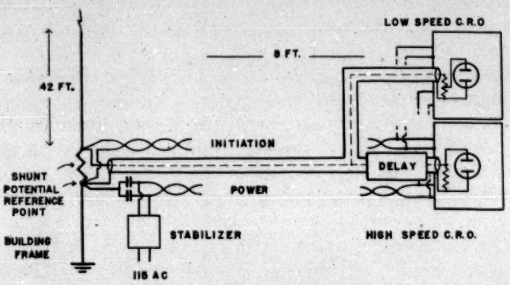


FIG. 3. External oscillograph connections designed to prevent disturbing surge currents in the potential measuring circuits. Indiscriminate "grounding" with resultant common impedances has been avoided, and all potentials referred to a single point in the circuit.

parts of the apparatus. The oscillograms must give a measurement of one voltage wave across a certain definite impedance and be completely insensitive to currents, voltages, and fields appearing elsewhere. Aside from the magnetic and electric field disturbances which are usually appreciated, the most disturbing element is unwanted surge current which flows from the point of measurement through the measuring leads to the oscillograph and leaves by various means. By electrical isolation of the oscillograph from all points except the shunt, this current can be made negligible. Fig. 3 shows the essential method of wiring. All electrical leads to the oscillograph come from a point at the same potential and follow adjacent paths so that the only surge current which

flows to the oscillograph is that required for operation and a negligible amount due to stray capacitance and slight voltage difference to the surroundings. This voltage difference has been kept a minimum by a short path from the shunt to the oscillograph.

THE BEAM CONTROL CIRCUIT.

The main circuit elements of the high-speed oscillograph comprise a high-voltage power supply, a circuit for controlling the initiation and duration of the cathode-ray beam, and a circuit for producing a sweep of the beam at a known rate.

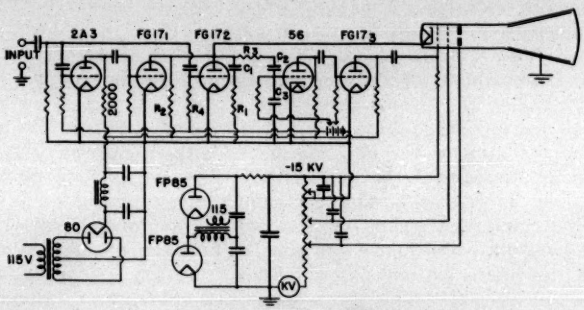


FIG. 4. Schematic circuit for the control of the beam of the high-speed cathode-ray oscillograph.

The design of the circuits for transient response has been greatly facilitated by the use of an oscillograph transient analyzer [22] by which means it is possible to determine surge performance of all circuits involved. Varieties of surges can be employed and traced through much in the manner of the ordinary steady-state oscillograph measurements. In the circuits described here some possible simplifications will be evident as no attempt has been made to carry the design beyond the reliable working state. For instance, it is feasible to operate both sweep control and beam control from the same circuit. Also, the use of suitable high-vacuum tubes in the trigger circuits in place of the Thyratrons would eliminate some variations in performance caused by thermal effects in

the mercury vapor and possibly permit more conventional control of the electrode potentials of the cathode-ray tube such as is common in television practice.

The beam circuit and power supply are shown in Fig. 4. A conventional voltage-doubler rectifier provides as much as 15 Kv. accelerating potential from cathode to screen for a tube of the 912 type. Usual operation is about 12 Kv. as this is sufficient for satisfactory recording. Ripple voltage must be held to a few volts at the divider. Proper operating potentials for the various electrodes are obtained from the divider shown. Superimposed upon the steady state values are the beam control voltages which turn the beam on for a sufficient time when the proper surge voltage is applied to the input.

The cathode-ray beam is turned on by the breakdown of the FG 17 Thyratrons either 1 or 2 shown in Fig. 4. The collapse of voltage across the Thyratrons on discharge lowers the potential of the cathode of the cathode-ray tube and thereby turns on the beam. With the proper circuit constants in the Thyratron discharge circuit which comprise mainly R_1, C_1, and R_2, the time during which the Thyratron conducts is controlled and hence the time of duration of the beam. With a limited current passed by the Thyratrons, deionization occurs after the time interval and the recovery of voltage removes the beam. Since the deionization time varies with different tubes and temperatures, the actual removal of the beam is accomplished by the discharge of Thyratron 3 which lowers the potential of the control grid of the cathode-ray tube. The correct delay of the discharge of this tube, and hence the beam duration time, is controlled by the circuit constants R_3, C_2, C_3 and the 56 tube acting as a phase inverter. Recovery of this Thyratron takes place in the same manner as the others. Deionization occurs because of limited current and the plate voltage recovers enabling the complete control circuit to be ready for a subsequent initiation.

A positive surge in the grid voltage of either Thyratrons 1 or 2 will cause a discharge. The method of obtaining this surge at the grids is of great importance since reliable initiation is essential at a known voltage and time related to the surge to be measured. When shunt voltage reaches a value

sufficient for recording by the oscillograph, positive surge voltage sufficient for initiation must appear on the grid of one of the Thyratrons. However, low rates of voltage change must be discriminated against, even though sufficient in magnitude to record.

The input terminals are connected directly to the lightning shunt so that the same voltage appears at this point that is applied through the delay cable to the deflection plates of the cathode-ray tube. If this voltage is of positive polarity with respect to the ground side of the shunt and oscillograph, it is applied directly to the grid of Thyratron 2 by means of capacitance coupling. The voltage which appears on the grid of 2, that is, the voltage across R_4, may be attenuated by the adjustment of R_4 and the capacitance leading to the input terminals. This attenuation is such that voltage waves which have a slow rate of change do not appear across R_4, and, by this means, discrimination is achieved, and records which are not suitable for this oscillograph are eliminated.

If the voltage appearing at the input terminals is of negative polarity, a modification in the arrangement is made for using this voltage to discharge a Thyratron. Many methods suggested themselves. The use of a transformer,[2,13] the application of negative voltage to the cathode,[2] or the reversal of the Thyratron were among these. As previously mentioned, the same discrimination against unwanted records had to be retained. Also, before any measurements were made, it was felt that as near ideal circuits as possible should be used. It was found in the laboratory that any wave within recording limits could be reversed in polarity by a low-impedance triode operating as a phase inverter in a resistance-coupled circuit. This is accomplished by the circuit of Fig. 4 where the negative surge is applied to the grid of a type 2A3 tube which has a low-resistance load in the plate circuit in addition to the variable load of the grid of Thyratron 1. In the plate circuit of the 2A3, a replica of the input wave is obtained but with polarity reversed. By the use of the low-resistance load and high-plate current, the distortion of the wave is negligible until the load imposed by the grid conduction of the Thyratron becomes a factor. This load must be supplied by the 2A3 or the time lag of the

Thyratron may become excessively long. In order for a Thyratron of the types used to conduct in less than a microsecond after voltage application to the grid, the amount of energy supplied to the grid must exceed a certain value.

AUTOMATIC SWEEP CIRCUIT.

Synchronized in time with the beginning of the beam is a voltage applied to one set of plates in the cathode-ray tube. This voltage changes at a known rate and sweeps the beam across the fluorescent screen. The sweep circuit is shown schematically in Fig. 5. As with the beam circuit, the input

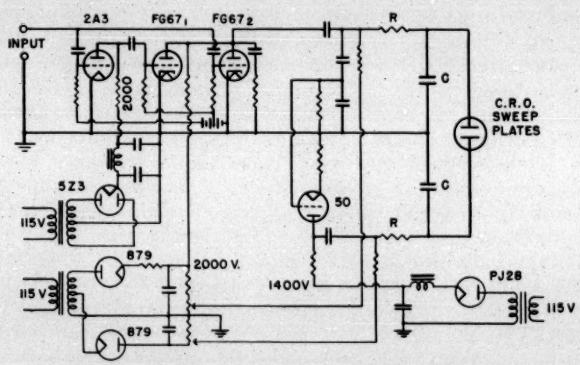


FIG. 5. Automatic sweep circuit of the high-speed oscillograph.

terminals are directly connected to the shunt. A positive surge causes the discharge of Thyratron 2, and a negative surge by means of the 2A3 phase inverter discharges Thyratron 1. The collapse of voltage is applied to the sweep plates through a resistance-capacitance circuit which gives an approximately logarithmic voltage on the sweep plates. In order that a balanced-to-ground sweep voltage be used for improved focus of the beam at various points on the screen, a type 50 tube used as a phase inverter supplies a balancing component to one of the plates which would otherwise be at ground potential.

The use of a non-linear sweep enables the first few microseconds of the wave to be well resolved, yet record the wave for a time of the order of 100 microseconds. However, a non-uniform sweep presents difficulties which, if overlooked, may lead to large errors. When such a sweep is used, the time scale which is subsequently applied to the oscillogram must be accurately matched by some means at zero time. Small shifts in the scale may lead to large errors in time.

THE LOW-VOLTAGE OSCILLOGRAPH.

Lightning current may continue through the shunt for a good fraction of a second without sufficiently rapid variations for high-speed recording. This fairly steady current must be

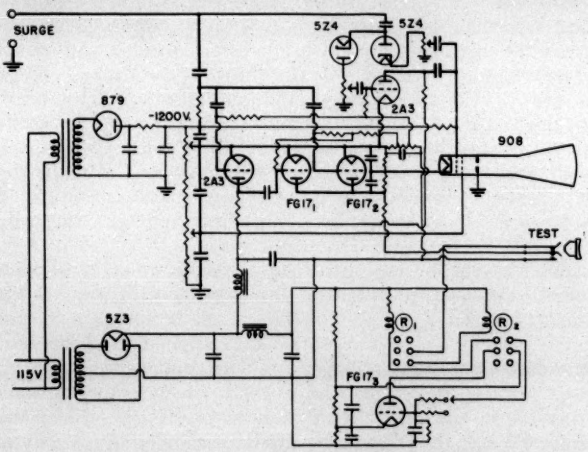


FIG. 6. Control circuit of the low-speed automatic oscillograph.

recorded on an instrument where the time scale extends over the complete time of the stroke. A fast sweep is not needed, but a record of considerable length is necessary. This sweep is supplied by the movement of film on the rotating drum, deflection being produced by the cathode-ray tube at right angles to film motion. Since the currents of long duration

are of lower magnitude, an oscillograph for this purpose may have a higher sensitivity with less writing speed than the high-speed oscillograph. The type 908 tube operated at 1200 volts serves this purpose. The design of this oscillograph was based to a large extent upon information obtained with a galvanometer-type instrument known as the Crater Lamp Oscillograph.[1,3,11 23,24]

The control circuit for the beam is shown in Fig. 6. Initiation of the beam is accomplished in the same manner as for the high-speed oscillograph. However, for this oscillograph the beam remains on for a length of time sufficient to record current in a stroke of very long duration. Also, no discrimination against low rates of voltage change is used. Initiation occurs when current exceeds thirty to fifty amperes and recording continues for a fixed time, while during this time the high-speed oscillograph may initiate and record many times, depending upon the nature of the stroke. Either a positive or negative surge produces initiation of the beam. In this circuit, however, the current drawn by the Thyratron that breaks down actuates the relay R_2 and is sufficiently great that the discharge is not self-extinguishing. Conduction is maintained until plate voltage is removed. Removal of voltage is accomplished by a time delay circuit comprising relays R_1 and R_2 and Thyratron No. 3. When relay R_2 is excited at the beginning of the stroke, contacts are closed which supply plate potential and positive grid potential to the Thyratron No. 3. This Thyratron is biased such that breakdown does not occur on the application of plate potential, but does so only after the grid potential changes sufficiently in a positive direction. Adjustment of the charging time constant in the grid circuit permits precise control of this time. When the Thyratron discharges, relay R_1 is excited momentarily and opens the power supply to relay R_2 and the initiation circuit Thyratrons. This removes the beam and ends the cycle of operation. For the purpose of setting the beam at the proper intensity and focus, a push button is provided which when pressed by the operator performs the same function as the breakdown of the initiation Thyratrons.

Since the beam may stay on for as long as a second and at times without much motion on the screen, it is desirable

to have some form of automatic intensity variation. Without such a control, a very wide trace is obtained on the moving film when the intensity is adjusted to the maximum and the velocity of the beam on the screen is low. If the intensity is reduced, then loss of recording speed occurs for high velocities. To provide this variation, a circuit comprising two 5Z4 rectifiers and a 2A3 used as an inverter applies a more positive potential to the control grid of the cathode-ray tube during portions of the stroke when the beam velocity is high. The voltage applied is proportional to the rate of change of voltage of the surge and is of the same polarity regardless of surge polarity. With the proper adjustments of magnitude and time constants, this circuit enables the tube to operate at a much higher intensity during high-beam velocity, thus providing improved recording of crest currents.

THE PHOTOGRAPHIC RECORDING SYSTEMS.

Photographic film is carried for each oscillograph in identical drums 10 inches wide and 29 inches in diameter as shown in Fig. 7 and Fig. 8. The surface of each drum is machined to close tolerance in the radius in order to permit the film to be used with a lens of small depth of focus. Each drum rides on a carriage which slides along a track in a light-tight metal housing. Two motors are provided for each drum. One motor rotates the drum by means of a belt drive. This motor operates continuously and rotates the drum one revolution in 0.7 seconds. The other motor drives the carriage along the track by means of a screw. This motor operates only after a lightning stroke has caused initiation and stops after translating the drum to a new exposure of film. Thus, superposition of records of separate strokes is prevented. With the 10-inch wide drum 10 exposures are provided so that as many strokes may be recorded without attention by an operator. The duration time of translation is controlled by a time delay relay which is excited either from contacts on the relay R_1 for the slow-speed oscillograph or an auxiliary relay circuit connected to the high-voltage sweep circuit. The relay circuits are so designed that initiation of one oscillograph produces only the translation of the proper drum. Often the low-voltage

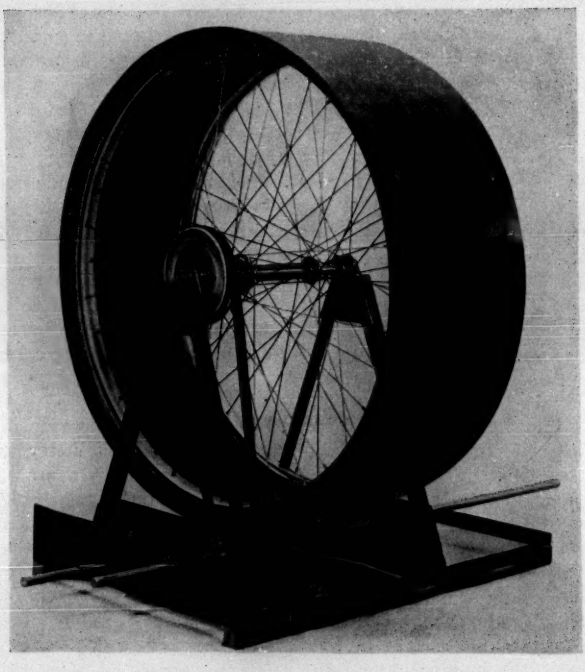


FIG. 7. Film drum and carriage with housing removed showing driving rod for supplying rotation, translation screw and carriage track.

oscillograph trips and records without any operation of the high-speed oscillograph. This occurs when the stroke consists of only the low-current component of long duration without superimposed current peaks.

The cathode-ray tube in each oscillograph is placed directly under the drum axis. A sleeve mounted in the floor of the drum housing carries a sliding cylinder which mounts the lens. Adjustment of focus is accomplished from the inside of the housing by viewing the image on the equivalent

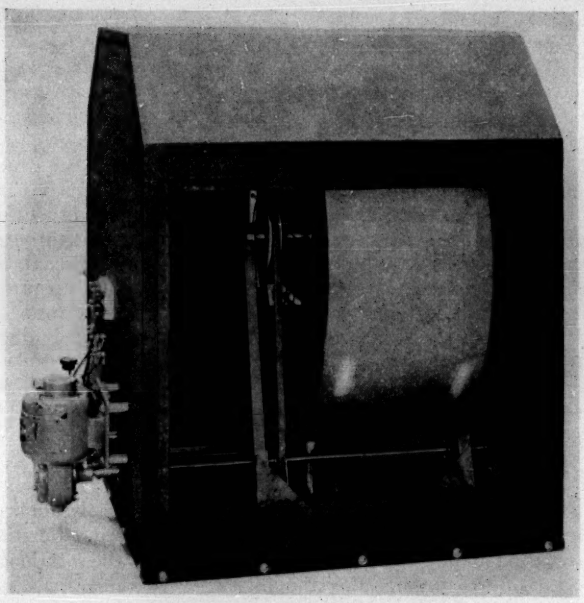


FIG. 8. Film drum housing with rear panel removed showing motor drive. Translation motor is obscured by the rotation motor. Sleeve for mounting the photographic lens is shown below the drum.

of a ground glass in the surface of the drum and adjusting the cylinder.

In order to view the tube screen and check operating conditions from time to time while the drums are loaded with film, each oscillograph has a viewing mechanism which consists of a mirror and shutter arrangement. The turning of a lever closes a shutter over the lens and then swings a mirror at 45 degrees between the screen and lens. The tube screen may be plainly viewed through the mirror and a spring return insures against the shutter remaining closed.

The high-speed oscillograph uses a lens of F/2 aperture

and 45 mm. focal length. The image is reduced about 10 times which gives a working aperture of F/2.2. The low-speed oscillograph uses an F/1.5, 50 mm. lens with a reduction of 1.7 times. The screens in both tubes are the short-persistence blue type which are photographed by high-speed film.* Film is obtained in rolls of the proper length and loaded in the dark. A single clamping bar across the drum surface holds the film securely.

In order to test proper operation of all circuits involved, provision is made for the application of surge current through the shunt. Capacitors charged above 5 Kv. are discharged through the shunt by means of a solenoid switch operated from a push button. A reversing switch permits tests of either polarity. The current produced through the shunt is a damped oscillation with an initial rate of rise to 10,000 amperes in a time of the order of 5 microseconds. This current is sufficient to cause initiation and be recorded by both oscillographs.

CALIBRATION OF RECORDS.

Each oscillogram must have two scales accurately applied by which the current and the time are determined. For the low-speed oscillograph, the time scale is fixed by the rotational speed of the drum. For the high-speed oscillograph, the output from a 200 KC. oscillator can be switched to the deflection plates and a record made. From this record a time scale is determined which is applied to any oscillogram obtained with the same accelerating voltage.

The current scale requires a determination of the current-voltage characteristic of the shunt, the voltage sensitivity of the cathode-ray tubes, and attenuation by any connecting networks between the shunt and the deflection plates. Shunt calibration was determined in the laboratory by means of metallic shunts in series with the lightning shunt. A high-current impulse generator permitted calibration to 150,000 amperes. The same shunt was used for several years and has been described.[2] The new lightning rod permitted a shunt with less restrictions. The shunt used with the new

* Agfa Superplenachrome Press.

lightning rod consists of a metallic element of 0.030 ohms and a non-linear element of Thyrite [2] in parallel as shown in Fig. 9. The metal shunt consists of a resistance ribbon folded over a mica card and securely clamped by insulating plates. Copper current terminals brazed to the ribbon extend at right angles and permit bolting to heavy copper

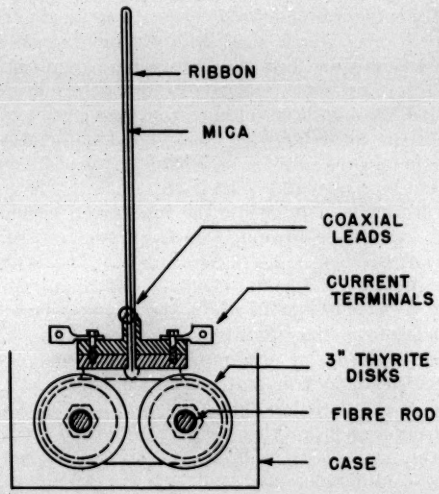


FIG. 9. Arrangement of the current conducting elements of the lightning shunt. Front cross section.

blocks which are the terminals for the non-linear assembly. Copper lugs which conduct the lightning current from the cable through the shunt to ground are also bolted at the same point. Potential leads in the form of a coaxial cable are connected at the junction of the ribbon and the copper current terminals. For insulation purposes the Thyrite assembly is liquid immersed. The use of a non-linear resistance material provides compression of the current scale so that reasonable accuracy can be obtained for a wide range of currents.

The voltage sensitivity of the high-speed oscillograph is determined directly upon the film as is the time scale. By means of switch positions the deflection plates may be short-circuited or may be excited from a known voltage from a d-c. source. Initiation of the oscillograph by means of the test set for the two switch positions provides on the film a zero potential line and a steady potential line from which the sensitivity is determined. For this calibration the film drum must not rotate. Such a calibration is placed on each film before its removal. For the low-voltage oscillograph both the sweep speed and sensitivity are determined by the application of a known voltage from the sixty-cycle power supply. This need only be done occasionally as operating voltages remain the same. All a-c. voltage to the equipment is provided by a constant voltage supply.

The deflection voltage for the high-speed oscillograph is obtained from the lightning shunt through a delay cable which retards the wave 1.9 microseconds. This coaxial cable is terminated in a surge impedance of 118 ohms as indicated in Fig. 3. Deflection voltage for the low-speed oscillograph has no delay from the shunt. However, a voltage divider is employed consisting of 200 ohms with a 100-ohm tap connected to the deflection plates and across this tap a non-linear resistor. This combination provides an additional voltage compression for this oscillograph. For low values of current through the main shunt the non-linear element has negligible conductance and the voltage on the deflection plates is approximately half that at the lightning shunt. However, with as much as 650 volts provided from the shunt only 100 volts appears on the plates. This compression enables the low-speed tube to cover a wider range of current values with better recording because of the reduction of beam velocities for high-current surges.

CHARACTERISTIC MEASUREMENTS.

Some of the first high-speed oscillograms [25] of lightning currents are shown in Fig. 10. These were obtained before the use of the type 50 tube and the balanced sweep voltage shown in Fig. 5; also before the beam removal circuit was employed which has since prevented the beam from having

occasional gaps as shown in these oscillograms. During 1938, 1939, and 1940, 30 strokes to the lightning rod have yielded some 60 high-speed oscillograms, each corresponding to a high-current surge or peak. In some strokes as many as 12 of these high-current peaks have been recorded while for others no peaks have occurred and only the low-speed oscillograph has given the records of the strokes. During these

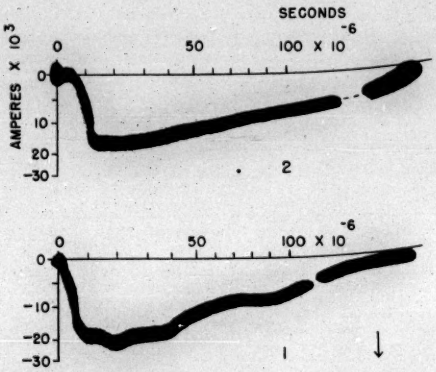


FIG. 10. High-speed oscillograms of two current peaks in a lightning stroke to the Empire State Building on Aug. 8, 1938, 6:57 P.M.; time between peaks 0.089 second. Arrow indicates direction of film movement.

years of operation it is felt from many factors involved, including accurate records of lightning activity, that every current peak of importance has been recorded and measured. Two heavy surges were recently obtained separated only 0.00092 seconds in time, which is less than the 1000 μs. repeating time for which the circuits were designed. Most high-speed oscillograms show a rapid initial rate of current rise and a slower rate of decay which is somewhat exponential in character as can be noticed. This same characteristic damped discharge occurs in the low-speed records where lower values and longer times are involved. Since the total length of a record is 90 inches and deflection is less than one inch, it is possible to reproduce only portions of the records

without loss of detail. Typical portions from separate strokes are shown in Fig. 11.

In addition to giving information on the nature of the lightning stroke and mechanism involved, such measurements tell much about disturbances that lightning can produce and the protection which is necessary. The form of the wave, the rate of change of current, the frequency of occurrence, and the duration of surges are of vital importance in the design of electrical equipment exposed to lightning. In conclusion, it should be pointed out that the physical size of the Empire State Building may possibly influence the char-

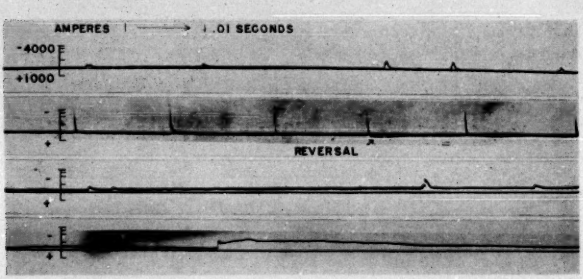


FIG. 11. Low-speed oscillograms showing peaks and continuing current. Appearing from the top down are portions of strokes of June 26, 1940, 1:38½ P.M.; 1:41 P.M.; July 27, 1940, 9:15 P.M.; and June 9, 1940, 7:04 P.M.

acter of the lightning discharge, by changing either the time of the beginning of the stroke, or the circuit constants of the discharge. Clouds which approach may be discharged prematurely. Upon being struck the building may be excited in the manner of a quarter wave antenna, or be considered to have reflections from the ground. The exact performance is difficult of determination mainly because it involves the electrical coupling of the stroke channel terminating at the building top. Following the initial current surge of the stroke approximately 2.5 μs. is required for the return of current by reflection from the ground. With increased sweep speeds used in 1940, many oscillograms have shown a similarity on the crest of the wave which may be related to this mechanism.[26]

Photographic evidence indicates the peculiarity of upward leaders [1] but any significant current distinction from lightning strokes to low objects remains to be seen.

Much effort and many ideas accumulated over a considerable time by associates in this problem have laid the foundation for this work. Many discussions with Mr. J. H. Hagenguth of this laboratory and Dr. K. B. McEachron, director of the lightning investigation, have been vital. Mr. C. J. Kettler of the Works Laboratory has carried out the construction details of the film housing and some other essentials. Much of the success is attributed to the careful operation by Mr. J. C. Russ, Mr. W. H. Eason, and Mr. P. C. Edwards during the electrical storms in New York City, in 1938, 1939, and 1940.

BIBLIOGRAPHY.

1. K. B. McEACHRON, "Lightning to the Empire State Building," JOURNAL OF THE FRANKLIN INSTITUTE, **227**, No. 2, 149 (1939).
2. J. H. HAGENGUTH, "Lightning Recording Instruments, Part 1," *G. E. Review*, **43**, 195 (1940).
3. J. H. HAGENGUTH, "Lightning Recording Instruments, Part 2," *G. E. Review*, **43**, 248 (1940).
4. EDGAR BELL AND A. L. PRICE, "Lightning Investigation on the 220 Kv. System of the Pennsylvania Power and Light Co. (1930)," *Trans. A. I. E. E.*, **50**, 1101 (1931).
5. R. R. PITTMAN AND J. J. TOROK, "Lightning Investigation on a Wood Pole Transmission Line," *Trans. A. I. E. E.*, **50**, 568 (1931).
6. HAROLD NORINDER, "On the Nature of Lightning Discharges," JOURNAL OF THE FRANKLIN INSTITUTE, **218**, No. 6, 717 (1934).
7. HAROLD NORINDER, "Lightning Currents and Their Variation," JOURNAL OF THE FRANKLIN INSTITUTE, **220**, No. 2, 69 (1935).
8. E. J. WORKMAN AND R. E. HOLZER, "Recording Generating Voltmeter for Study of Atmospheric Electricity," *Review of Scientific Instruments*, **10**, 160 (1939).
9. B. F. J. SCHONLAND AND H. COLLENS, "Progressive Lightning, I," *Proceedings of the Royal Society*, **143**, Series A, No. 849, 654 (1934).
10. E. J. WORKMAN, J. W. BEAMS, AND L. B. SNODDY, "Photographic Study of Lightning," *Physics*, **7**, 375 (1936).
11. K. B. McEACHRON, "Multiple Lightning Strokes," *Trans. A. I. E. E.*, **53**, 1633 (1934).
12. O. ACKERMANN, "A Cathode-Ray Oscillograph with Norinder Relay," *Trans. A. I. E. E.*, **49**, 467 (1930).
13. R. H. GEORGE, "A New Type of Hot Cathode Oscillograph and Its Application to the Automatic Recording of Lightning and Switching Surges," *Trans. A. I. E. E.*, **48**, 884 (1929).

14. HAROLD NORINDER, "The Cathode Oscillograph as Used in the Study of Lightning and Other Surges on Transmission Lines," *Trans. A. I. E. E.*, **47**, 446 (1928).
15. H. P. KUEHNI AND S. RAMO, "A New High-Speed Cathode-Ray Oscillograph," *Trans. A. I. E. E.*, **56**, 721 (1937).
16. P. L. BELLASCHI, "Lightning Currents in Field and Laboratory," *Elec. Engg.*, **54**, 837 (1935).
17. J. F. PETERS, "The Klydonograph," *Electric World*, **83**, 769 (1924).
18. C. M. FOUST AND H. P. KUEHNI, "The Surge Crest Ammeter," *G. E. Review*, **35**, 644 (1932).
19. W. W. LEWIS AND C. M. FOUST, "Lightning Investigations on Transmission Lines, IV," *Trans. A. I. E. E.*, **53**, 1180 (1934).
20. H. GRÜNEWALD, "Die Messung von Blitzstromstarken und Freileitungsmasten," *Electro-Technische Zeitschrift*, **55**, 505 (1934).
21. C. F. WAGNER AND G. D. MCCANN, "New Instruments for Recording Lightning Currents," *Trans. A. I. E. E.*, **59**, 1061 (1940).
22. N. ROHATS, "An Oscillograph Electric Transient Analyzer," *G. E. Review*, **39**, 146 (1936).
23. W. A. MCMORRIS, M. A. RUSHER, AND J. H. HAGENGUTH, "The Crater Lamp Oscillograph," *G. E. Review*, **37**, 514 (1934).
24. P. SPORN AND I. W. GROSS, "Expulsion Protective Gaps on 132 Kv. Lines," *Trans. A. I. E. E.*, **54**, 56 (1935).
25. K. B. MCEACHRON, "Wave Shapes of Successive Lightning Current Peaks," *Electrical World*, **113**, No. 6, 428 (1940).
26. K. B. MCEACHRON, "Lightning to the Empire State Building," *Trans. A. I. E. E.*, **60**, 885 (1941).

CONTROL PROBLEMS.

BY

NICHOLAS MINORSKY, Ph.D.; E.E.

CONTENTS.

1. Introduction 451

Part I.

Application of Operators to Linear Control Problems.

Page

2. Elements of controlled system; degrees of freedom 453
3. Differential control operator 456
4. Distributive rule of operators; resultant operator 458
5. Integrated control terms in operational notations 459
6. Direct and indirect control operators; non commutativity of operators .. 459
7. Symbolic multiplication of operators by δ^{λ} 461
8. Differential equation of controlled process; effect of integrated terms in the operator 463
9. Examples of application of operators to control problems 464

Part II.

General Properties of Linear Systems Acted upon by Linear Controls.

10. Classification of control systems 466
11. Fundamental control system and its properties 467
12. Non-fundamental control systems; effect of residue operators on the controlled process 468
13. Principle of superposition; control terms 471
14. Control matrix and its reduction to direct control terms 472
15. Sheared matrix of control terms; equivalent control terms 474
16. Examples of equivalent control terms. Floating zero in a controlled process 475
17. Examples of application of operators to the analysis of various control terms. Controlled pendulum 479

INTRODUCTION.

1. In recent years the number of control problems of interest to industry and applied science, as well as to the National Defense has increased very appreciably. One of the reasons for that is due undoubtedly to the availability of new continuously acting control devices such as modern thermionic circuits and apparatus. Solution of problems which appeared to be either difficult or even impossible about two decades ago

or so, becomes feasible now, owing to the use of these new control means. Furthermore, a mathematical formulation of control problems with the use of circuits possessing a continuous control action becomes considerably simpler than in the past when discontinuously acting devices such as contactors, relays, etc., were the only ones available.

In view of these facts, there is a steadily growing tendency in the technical literature of recent years [1] (1) to consider the problems of control from a somewhat broader point of view, than in the past.

The present paper attempts to follow this general trend towards greater generalizations in this particular field. Parts I, II and III of these notes deal with problems of linear controls defined below. Operational notations are consistently used throughout the first three parts of this paper.

Although the use of linear operators in connection with problems having one degree of freedom to which this paper is limited, is not so indispensable as in problems with several degrees of freedom, the application of operational notations facilitates generalizations particularly in view of the considerable abbreviations of the mathematical text.

Classification of the different control systems described in this paper, is based on analytical features of operators characterizing the control process rather than on material details of corresponding control systems as is customarily done in the existing technical literature. Although such a procedure may appear somewhat pedantic, particularly in view of the practical significance of the subject, this form of classification is believed advantageous in that it correlates more definitely the properties of different control systems with those of the differential equations governing the controlled process.

Following this line of thought control problems will be divided into linear and non-linear ones.

The first three parts of this paper are devoted to a study of linear problems; to this class belong practically all existing control systems.

Part IV deals with non-linear control systems. Although the intrinsic difficulties of this particular field have retarded considerably the advent of these non-linear methods, it is

[1] Figures in text relate to the list of references.

thought that its potential possibilities may be of a sufficient interest to warrant at least a rather superficial review here.

It is hoped that this paper, following, so to say, "a middle course" between the purely practical problems of control and the corresponding mathematical generalizations may be helpful to engineers interested in the broader aspects of the subject, as well as to mathematical physicists. The latter will find in this field a series of interesting theoretical problems concerning definite engineering applications.

PART I.

Application of Operators to Linear Control Problems.

2. Elements of a Controlled System: Degrees of Freedom.

In any system subject to a control, there are three principal elements to be considered:

A. System (*A*) to be controlled per se.
B. Control system (*S*).
C. External disturbing action (*D*).

One can represent different cases symbolically as follows:

1. $(A) = 0$—the free "motion" (or, more generally, evolution) of the system (*A*) in the course of time.
2. $(A) = (D)$—the behaviour of the "natural" (i.e. uncontrolled) system (*A*) under the effect of a given disturbance (*D*).
3. $(A + S) = 0$—the behaviour of a controlled system in absence of any external disturbance.
4. $(A + S) = (D)$—the behaviour of a controlled system when acted upon by a disturbance.

These symbolic notations are in reality differential equations characterizing the evolution of either uncontrolled (1, 2) or controlled (3, 4) process in the course of time.

A great majority of controlled systems are systems with one degree of freedom, i.e. systems describable in terms of one single dependent variable θ which is the deviation, or departure, from a pre-determined "zero point" around which the control mechanism is supposed to maintain the system (*A*). Problems of temperature, flow, level and speed control are

obviously problems with one degree of freedom inasmuch as the control process is generally describable in terms of one single dependent variable θ.

In problems of dynamical control the situation is somewhat more complicated, the degree of complexity depending on the particular problem; very frequently even in the same problem, conditions may differ as the mode of control is changed. For example, in problems in which the control action S is derived from forces (or torques) *extraneous* to the system (A), a dynamical problem can still be considered as having one single degree of freedom; on the other hand, in cases in which the control action is obtained from forces partially extraneous to, and partially internal in the system (A), the problem is more complicated.

Consider, for instance, the problem of antirolling stabilization of ships. The same problem may present itself in two different aspects according to the nature of stabilizing moment S adopted. If one considers for example, the stabilization by fins arranged on both sides of the vessel and operated by a control mechanism so as to oppose by their "lifts" the disturbing moment caused by waves—the problem is clearly that with one degree of freedom. The stabilizing action S in this case, comes from an extraneous source—the relative impingement of water particles on the fins. The counterpart of this reaction—the "drag"—is not felt in this particular "rolling degree of freedom" and merely reduces the ship's speed. Everything happens as if the stabilizing moment S due to the "lift" of fins were the only extraneous moment applied to system (A)—the ship—insofar as the reaction—the "drag"—is relegated to a degree of freedom (forward speed) in which it has no direct relation to the control process. Thus, the latter can be expressed in terms of variables relative to one single degree of freedom θ—angle of roll—and can be represented symbolically by $(A + S)$, the assumption being made that (S) is related to $\theta, \dot{\theta}, \ddot{\theta}$ by some kind of a functional relation, which constitutes *the control problem* proper.

Consider now the same problem but with a different method of producing the stabilizing moment, e.g. displacement of water ballast. The system (A) to be controlled is now more complicated insofar as it comprises two systems

(A_1)—the ship, and (A_2)—the ballast. Even without any control at all, the combination $(A_1 + A_2)$ forms a "coupled system" (the so-called, Frahm's, or passive, tanks); the system $(A_1 + A_2)$ is more complicated than either (A_1) or (A_2) considered separately; in fact, if (A_1) or (A_2) are individually describable in terms of differential equations of the second order, the combination $(A_1 + A_2)$ is governed by a differential equation of the 4th order as follows from the theory of passive tanks (2). If, in addition to this, the motion of the ballast is controlled by a local source of power (air pressure or pumps), the control action is generally applied to (A_2) but, by virtue of internal couplings between (A_1) and (A_2) is felt on (A_1) as well; furthermore, the reaction from (A_2) to (A_1)—e.g. the axial thrust of the pumps applied to the ship—is not shifted in this case to a degree of freedom where it is of no further influence on the control process, but, generally operates in the same "rolling degree of freedom" (3, 4).

It is seen, thus, that the problem in this case is considerably more complicated than in the first case, owing to the existence of two degrees of freedom: θ, the angle of roll and φ—the angle of levels in the tanks with respect to their middle position.

The problem of control in this case consists in establishing a predetermined relation between θ, $\dot{\theta}$, $\ddot{\theta}$ on one hand, and φ, $\dot{\varphi}$, $\ddot{\varphi}$ on the other hand, that is—between the ship's motion and the ballast's motion.

This reduces the problem from two degrees of freedom to one single degree of freedom in view of the relation between θ and φ, established by control.

It must be noted, however, that such a reduction generally lowers the order of the differential equation of the controlled process. Thus, in the above example, the differential equation of the passive system $(A_1 + A_2)$ is generally of the fourth order. The process of "activation" above considered reduces the order of the differential equation of the system to the third order in view of the fact that the control provides an additional relation between θ and φ (5).

These complications are merely mentioned here without elaborating their theory any further, as an example of diffi-

culties arising in more complicated cases of dynamical controls.

In the following we shall be concerned only with systems having one degree of freedom. This, however, does not limit the generality of conclusions so obtained.

3. *Differential Control Operators.*

In this Part I behavior of linear systems (A) controlled by linear controls, that is, by such control systems (S) the analytical features of which can be represented by a linear assembly of individual "control terms"—is investigated. In view of the fact that control problems deal generally with motions (or evolution) of the system (A) in the vicinity of an equilibrium point, simplifications due to the Method of Small Motions are generally available. As is known, these simplifications consist in expanding different functions appearing in the problem by Taylor's formula and in neglecting terms of higher than first order. This permits reducing the differential equation of a more general form to a differential equation with constant coefficients, which, for simple dynamical problems, is generally of the second order; for problems of control of temperature, pressure, flow, etc., the differential equation is usually of the first order. Finally, for more complicated dynamical "coupled" systems, it is generally of the fourth order.

In the following, use will be made of the operational notations particularly applicable to linear problems. A few remarks in connection with the use of operators may be useful for the following.

Assume, for example, that the motion (or evolution) of the original uncontrolled system (A) be governed by a simple differential equation of a damped pendulum type:

$$A_2\ddot{\theta} + A_1\dot{\theta} + A_0\theta = D, \qquad (1)$$

in which θ is the dependent variable (e.g. angle of departure from equilibrium); A_2, A_1, A_0 are certain constant coefficients characterizing system (A) and D is the external disturbing (or "forcing") function, which is generally a function of the independent variable t-time. The dependent variable θ varies

with different problems. For example in dynamical control problems (e.g., stabilization of ships, airships, automatic steering, etc.), θ generally means the angular deviation of (A) from its predetermined value θ_0 which the control mechanism is supposed to maintain. In other problems of control (e.g., temperature, pressure, etc.), θ also designates a corresponding departure of the basic variable of the problem from its equilibrium value θ_0.

Instead of dealing with differential equation (1) it is frequently convenient to describe it by means of *an operator* $[A]$ defined as follows:

$$[A] = \left[A_2 \frac{d^2}{dt^2} + A_1 \frac{d}{dt} + A_0 \right] = [A_2\delta^2 + A_1\delta + A_0], \qquad (2)$$

in which symbols $\delta = \frac{d}{dt}$; $\delta^2 = \frac{d^2}{dt^2} \cdots$ are used; the differential equation (1) is then obtained by applying operator (2) to the variable θ

$$[A]\theta = [A_2\delta^2 + A_1\delta + A_0]\theta = A_2\ddot{\theta} + A_1\dot{\theta} + A_0\theta = D, \qquad (3)$$

which gives the original differential equation (1). So far there is no particular advantage in substituting one notation for the other. The advantages will, however, appear when we extend these operational notations to control problems.

Assume now that in addition to the intrinsic parameters of the system (A) viz. A_2, A_1, A_0 which appear in equation (1) and determine the behavior of system (A), per se, we add a *control action* S which will be subordinated to certain conditions which we can select at will. For the controlled system, the differential equation will then be:

$$A_2\ddot{\theta} + A_1\dot{\theta} + A_0\theta + S = D. \qquad (4)$$

In order to specify the matter further, assume that the problem in question is a dynamical one, e.g. of a damped pendulum type.[2]

[2] Although the following is devoted mainly to the analysis of behaviour of dynamical control problems, the general methods outlined below are applicable to other problems of control (e.g. temperature, speed, flow, pressure etc.). In each particular problem it is necessary to establish the form of operator $[A]$ to begin with.

The variable θ is then the angular departure. Since the physical dimensions of $A_2\ddot{\theta}$, $A_1\dot{\theta}$, $A_0\theta$ and D are clearly moments or torques, S must also be of the same dimension—*the control moment*. The control problem becomes definite if one specifies the analytical form of the control moment S. Here, at least theoretically, one comes across an infinite variety of possibilities. Consider the simplest case, viz.: S is expressed by the same analytical form as the differential equation (1) itself, i.e.

$$S = a_2\ddot{\theta} + a_1\dot{\theta} + a_0\theta. \tag{5}$$

Incidentally, it is worthwhile mentioning here, that this simple form for S which suggests itself from the form of differential equation (1), is also the most important one for control purposes as will appear from the following. Equation (5) is not a differential equation but a *differential form* of the control action S.

But since the differential operators are used for a designation of differential forms, it is apparent that (5) can be represented by a differential operator which we can write[3] as

$$[S] = [a_2\delta^2 + a_1\delta + a_0], \tag{6}$$

using the same notations as before.

The differential operator $[S]$ is the *control operator of the problem*. If we apply now operator $[S]$ to the independent variable θ we have:

$$[S]\theta = a_2\ddot{\theta} + a_1\dot{\theta} + a_0\theta,$$

we obtain thus equation (5) of definition for S with which we started.

4. *Distributive Law of Operators—Resultant Operator* $[R]$.

Linear operators follow the distributive law, i.e.

$$[A] + [S] = [A + S], \tag{7}$$

that is, the sum of operators $[A] + [S]$ is equal to the operator $[A + S]$ of the sum. This is self-evident from the above definitions. Whence the rule: in order to find the

[3] In the following, unless specified to the contrary, square brackets will designate operators. Thus $[A]$ will mean "operator A"; $[S]$—"operator S" etc.

differential equation of the controlled system $(A + S)$ it is sufficient to form the *resultant operator* $[R]$ by combining operators $[A]$ and $[S]$ and to operate with $[R]$ so formed on θ—which gives the differential equation of the controlled process.

5. *Extension of Operational Notations to Integrated Control Terms.*

The control operator $[S] = [a_2\delta^2 + a_1\delta + a_0]$ so far considered, is as yet a very special control operator. The most general form of a linear control operator $[S]$ is:

$$[S] = [\cdots + a_3\delta^3 + a_2\delta^2 + a_1\delta^1 + a_0 + a_{-1}\delta^{-1} + a_{-2}\delta^{-2} + \cdots], \quad (8)$$

that is,

$$[S] = [\cdots + S_3 + S_2 + S_1 + S_0 + S_{-1} + S_{-2} + \cdots],$$

where

$$[S_3] = [a_3\delta^3], \quad \cdots [S_0] = [a_0], \quad \cdots [S_{-2}] = [a_{-2}\delta^{-2}] \cdots.$$

In this general form, we come across the terms $a_{-1}\delta^{-1}$, $a_{-2}\delta^{-2}\cdots$ containing negative powers of δ such as: δ^{-1}, $\delta^{-2}\cdots$. The significance of these negative powers is obvious. In fact if $\delta = \frac{d}{dt}$ means a differentiation, clearly δ^{-1}, designating an inverse process, means an *integration*. We are thus led to identify operators $[S_{-1}] = [a_{-1}\delta^{-1}] = [a_{-1} \int dt]$; $[S_{-2}] = [a_{-2}\delta^{-2}] = [a_{-2} \int dt \int dt]$ etc. One can now condense the explicit expression (8) of a linear control operator $[S]$ in a form:

$$[S] = \left[\sum_{i=-\lambda}^{i=+\mu} a_i\delta^i \right], \quad (9)$$

which includes both positive powers of δ (i.e. differentiations) as well as the negative ones (i.e. integrations).

6. *Direct and Indirect Control Operators. Non-Commutativity of Operators.*

Control operators of the form (9) will be called *direct control operators*. For such operators, the control action S is

represented *directly* by a linear assembly of terms $\sum_i a_i\delta^i$ characterizing the individual control terms. Operator (6), previously cited, is an example of a direct operator, characterizing the form (5) of control.

In applications the control action S, instead of being controlled *directly* as in the above example (5), is frequently controlled *indirectly* either through its rates $\frac{dS}{dt}, \frac{d^2S}{dt^2} \cdots$ or through its integrated values $\int Sdt, \int dt \int Sdt \cdots$. Operators corresponding to such indirect methods of control will be called *indirect control operators;* such indirect operators can be defined by a symbolic equation:

$$\delta^\lambda[S] = \sum_i [a_i\delta^i] \cdots. \tag{10}$$ [4]

For $\lambda = 0$ equation gives a direct control operator.

Operators, generally, do not follow the commutative law, that is:

$$[A][B] \neq [B][A].$$

This can be seen already from the consideration of elementary operators $[\delta]$ and $[\delta^{-1}]$. In fact, when operating on a dependent variable θ,

$$[\delta][\delta^{-1}]\theta = \frac{d}{dt}\int \theta dt = \theta$$

whereas

$$[\delta^{-1}][\delta]\theta = \int \frac{d\theta}{dt} dt = \theta + \theta_0,$$

where θ_0 is a constant of integration; the two operations are, thus, not quite identical. In the following, in order to abbreviate the text, the last mentioned operator $[\delta^{-1}][\delta]$ will be considered as commutative with $[\delta][\delta^{-1}]$ with that difference however that we shall write:

$$[\delta^{-1}][\delta]\theta = \theta_*, \tag{11}$$

the asterisk indicating that one constant of integration has thus appeared; for $\lambda = -2, -3 \cdots$ etc., the number of constants of integration is accordingly $2, 3 \cdots$.

[4] See section 7 below; λ in the following designates a positive integer.

7. *Symbolic Multiplication of a Control Operator* $[S]$ *by* δ^{λ}.

In the following, it will be necessary to transform the indirect control operators into the direct ones by considerations of dimensional homogeneity of terms of the differential equation. For this purpose it is necessary to consider the process of multiplication of an operator $[S]$ by δ^{λ}. It is sufficient to consider two cases when $\lambda = 1$ and $\lambda = -1$. Generalization for other cases does not present any difficulty.

Consider, for example, an indirect control operator:

$$[S^1] = \left[\frac{dS}{dt}\right] = [\delta S] = \delta[S] = [\cdots a_2\delta^2 + a_1\delta + a_0\delta^0 + a_{-1}\delta^{-1} + a_{-2}\delta^{-2} + \cdots]. \tag{12}$$

Multiplying it by δ^{-1} we have:

$$\delta^{-1}[S^1] = \delta^{-1}\delta[S] = [S_*] = [\cdots a_2\delta + a_1\delta^0 + a_0\delta^{-1} + a_{-1}\delta^{-2} + \cdots], \tag{13}$$

use being made of the previously established rule (11), so that one has:

$$\delta^{-1}[S^1] = \delta^{-1}\delta[S] = [S_*] = \sum_i a_i\delta^{(i-1)}. \tag{14}$$

More generally, for

$$[S^{\lambda}] = \delta^{\lambda}[S] = [\cdots a_2\delta^2 + a_1\delta + \cdots], \tag{15}$$

in which the control action is controlled through its λth rate, one finds upon multiplication with $\delta^{-\lambda}$:

$$\delta^{-\lambda}[S^{\lambda}] = \delta^{-\lambda}\delta^{\lambda}[S] = [S_*] = \sum_i a_i\delta^{(i-\lambda)} \cdots. \tag{16}$$

In this case one should attach λ asterisks to the reduced direct control operator so obtained. This will be designated still with one asterisk as $[S_*]$ in order not to over-burden these symbolic notations.

Likewise, when the control operator is of the form:

$$[S^{-1}] = \left[\int S dt\right] = \delta^{-1}[S] = [\cdots + a_2\delta^2 + a_1\delta + a_0\delta^0 + a_{-1}\delta^{-1} + \cdots], \tag{17}$$

that is, when the integrated quantity is controlled directly

and S itself only indirectly, a multiplication by δ gives

$$\delta[S^{-1}] = \delta\left[\int Sdt\right] = \delta\delta^{-1}[S] = [S] = \sum_i a_i\delta^{(i+1)} \quad (18)$$

and, similarly, for $[S^{-\lambda}]$ controlled directly:

$$[S^{-\lambda}] = \int dt \int dt \cdots \int Sdt = \delta^{-\lambda}[S]$$
$$= [\cdots + a_2\delta^2 + a_1\delta + \cdots],$$

a multiplication by δ^λ gives:

$$\delta^\lambda[S^{-\lambda}] = \delta^\lambda\delta^{-\lambda}[S] = [S] = \sum_i a_i\delta^{(i+\lambda)}. \quad (19)$$

In this case no asterisks should be attached to the reduced direct operator since no constants of integration appear in this case.

It is, thus, possible to pass from an indirect control operator to a direct one by multiplying the former either by $\delta^{-\lambda}$ or δ^λ according to whether the rate, or the integrated value of the control action is directly controlled.

In summing up, a multiplication of an operator by δ^λ follows the rules of algebra, on condition of attaching an asterisk to the reduced control operator when a multiplication by a negative power of δ is involved in the process of such reduction.

The following example gives an illustration of this process of reduction of operators to a direct form.

Consider again the stabilization problem and assume that the stabilizing moment S is controlled indirectly by its rate as follows:

$$\frac{dS}{dt} = a_2\ddot{\theta} + a_1\dot{\theta} + a_0\theta. \quad (20)$$

If the stabilization is obtained by shifting water ballast between the tanks, the physical significance of equation (20) of control is: the rate of flow of ballast is controlled in proportion to a linear combination of angular acceleration $\ddot{\theta}$, angular velocity $\dot{\theta}$ and angle θ of roll.

The control operator corresponding to equation (20) is:

$$[S^1] = \left[\frac{dS}{dt}\right] = \delta[S] = [a_2\delta^2 + a_1\delta + a_0]\cdots. \quad (21)$$

This is obviously an indirect control operator. Multiplying it by δ^{-1} one has:

$$\delta^{-1}[S^1] = \delta^{-1}\delta[S] = [S_*] = [a_2\delta + a_1 + a_0\delta^{-1}]. \quad (22)$$

Operator (22) is now a direct one.

Physical significance of constants of integration (i.e. asterisks in operational notations used) is explained in section 16 below.

8. *Differential Equation of Controlled Process. Effect of Integrated Terms in the Resultant Operator* $[R]$.

If the resultant operator $[R]$ is formed, the differential equation of the controlled process is $[R]\theta = D$, provided $[R]$ does not contain negative powers of δ i.e. integrated control terms.

In case $[R]$ does contain such terms $\delta^{-\lambda}$, λ being the largest number among negative exponents, the *whole differential equation including the right-hand term D, if any, must be multiplied by* δ^λ *i.e. differentiated* λ *times*, which gives:

$$\delta^\lambda[R] = [\delta^\lambda R] = \delta^\lambda D = \frac{d^\lambda D}{dt^\lambda}.$$

Operator $[\delta^\lambda R]$ now has no terms with negative powers of δ. However, the differential equation has as its right hand term $\frac{d^\lambda D}{dt^\lambda}$ and not D.

The most frequently encountered case in practical applications is when $\lambda = 1$ (equation 22). The differential equation of the controlled process is then $[\delta R]\theta = \delta D$. In other words, the presence of an integrated term $a_{-1}\delta^{-1}$ in the resultant operator $[R]$ raises the order of the original differential equation by one unit; furthermore, as "external disturbance" in such a case appears $\frac{dD}{dt}$ and not D itself.

The physical significance of such "integrated control terms" in the resultant operator is explained in section 17.

9. *Examples of Application of Operators to Control Problems.*

A few examples are given below for the purpose of illustrating the application of these operational notations to control problems.

(1) Assume again that a dynamical system (A) is characterized by operator: $[A] = [A_2\delta^2 + A_1\delta + A_0]$. Assume, further, that this system is controlled by a control moment S through its rate $\frac{dS}{dt}$ which is controlled in proportion to, and in phase with the angular acceleration $\ddot{\theta}$ of system (A).

The control operator here is, clearly:

$$[S^1] = \left[\frac{dS}{dt}\right] = \delta[S] = [a\delta^2].$$

Multiplying it by δ^{-1} we have: $\delta^{-1}\delta[S] = [S_*] = [a\delta]$. With the reduced direct control operator $[S_*]$ the resultant operator is:

$$[R] = [A + S_*] = [A_2\delta^2 + (A_1 + a)\delta + A_0] \cdots. \quad (23)$$

The differential equation of controlled motion is then:

$$[R]\theta = [A_2\delta^2 + (A_1 + a)\delta + A_0]\theta = D \cdots. \quad (24)$$

Written explicitely, it is:

$$A_2\ddot{\theta} + (A_1 + a)\dot{\theta} + A_0\theta = D. \quad (25)$$

Differential equation (25) gives the desired answer; it is seen that a control of this type modifies the natural damping A_1 of the system (A).

(2) Assume now that for the same system (A) the rate of the rate $\frac{d^2S}{dt^2}$ of the moment is controlled so as to be in phase with, and proportional to the angular velocity of the pendulum. The control operator is

$$[S^2] = \left[\frac{d^2S}{dt^2}\right] = \delta^2[S] = [a\delta].$$

Multiplying this by δ^{-2} we have the reduced direct control operator

$$\delta^{-2}\delta^2[S] = [S_*] = \delta^{-2}[a\delta] = [a\delta^{-1}].$$

The resultant operator is:

$$[R] = [A + S_*] = [A_2\delta^2 + A_1\delta + A_0 + a\delta^{-1}]. \quad (26)$$

In this form it contains a negative power of δ; multiplying the differential equation $[R]\theta = D$ by δ, we have

$$\delta[R]\theta = [\delta R]\theta = [A_2\delta^3 + A_1\delta^2 + A_0\delta + a]\theta = \delta D, \quad (27)$$

whence the differential equation of the control process

$$A_2\dddot{\theta} + A_1\ddot{\theta} + A_0\dot{\theta} + a\theta = \frac{dD}{dt} \cdots. \quad (28)$$

(3) Assume now that the integrated value of the moment $\int Sdt$ is so controlled as to be proportioned to, and in phase with the angular velocity of the pendulum. The control operator is

$$[S^{-1}] = \left[\int Sdt\right] = \delta^{-1}[S] = [a\delta] \cdots. \quad (29)$$

Multiplying this operator by δ we have the reduced direct control operator:

$$\delta\delta^{-1}[S] = [S] = \delta[a\delta] = [a\delta^2] \cdots. \quad (30)$$

Whence the resultant operator

$$[R] = [A + S] = [(A_2 + a)\delta^2 + A_1\delta + A_0]. \quad (31)$$

The differential equation of controlled motion will be then:

$$(A_2 + a)\ddot{\theta} + A_1\dot{\theta} + A_0\theta = D.$$

These few examples give a sufficient illustration of application of operators to the problems of control. A few additional features of operators in connection with control problems are indicated below.

PART II.

General Properties of Linear Systems Acted upon by Linear Controls.

In this Part II the general behavior of controlled systems is investigated. The study is limited to linear systems (A) actuated by linear controls i.e. such systems the differential operators of which are linear.

To this class belong, practically, all existing control systems with one degree of freedom; for such systems operator $[A]$ is generally of the second order for dynamical problems, and of the first order for problems of flow, temperature, pressure etc.

For more complicated dynamical problems having two degrees of freedom, such as system (A) of passive tanks, operator $[A]$ is generally of the fourth order in view of couplings provided between the component systems (A_1) and (A_2) of each degree of freedom. Action of control in such a case is equivalent to the establishment of an additional relation between the degrees of freedom (problem of activated tanks) which reduces the order of the resultant operator $[R]$ of the controlled system. The resultant system so obtained thus again becomes a system with one degree of freedom, while still being linear.

It is seen, thus, that investigation of linear systems of this kind is of a fundamental importance for a great variety of control problems. In Part II idealized conditions when no time lags exist in the control system, are investigated. The effect of a retarded control action in connection with the same linear problems is studied in Part III.

10. *Classification of Control Systems.*

We shall now consider the principal properties of different modes of linear controls, when applied to a linear system (A) to be controlled.

Consider again two operators: $[A]$ of the second order characterizing system (A) and the control operator $[S]$ characterizing the control system (S).

$$[A] = [A_2\delta^2 + A_1\delta + A_0]$$
$$[S] = [\cdots a_3\delta^3 + a_2\delta^2 + a_1\delta + a_0 + a_{-1}\delta^{-1}\cdots] = [S^{\cdot}] + [S_0] + [S_{\cdot}], \quad (32)$$

where

$$[S_0] = [a_2\delta^2 + a_1\delta + a_0]; \qquad [S^{\cdot}] = [\cdots + a_4\delta^4 + a_3\delta^3]$$
$$[S_{\cdot}] = [a_{-1}\delta^{-1} + a_{-2}\delta^{-2} + \cdots].$$

Equation (32) represents operator $[S]$, as consisting of three "component operators," viz.,

(1) $[S_0]$ which will be referred to as *fundamental component* of the control operator, or, simply, as *fundamental operator* and:

(2) $[S^{\cdot}]$ and $[S.]$ as "high order," and "low order residue operators" respectively.

With these definitions of component operators one can introduce the following classification of control systems, the significance of which is explained below.

A control system is said to be fundamental if $[S] = [S_0]$, that is, if *its operators* $[S]$ *is reduced only to its fundamental component* $[S_0]$ *so that the residue* operators $[S^{\cdot}] = [S.] = 0$. If either $[S^{\cdot}]$ or $[S.]$ or both, are not zero, the control system is a *non-fundamental* one.

11. *Fundamental Control Systems.*

Since for such systems $[S] = [S_0]$, the resultant operator is:

$$[R] = [A + S_0] = [(A_2 + a_2)\delta^2 + (A_1 + a_1)\delta + (A_0 + a_0)] \quad (33)$$

and the corresponding differential equation for the controlled system $(A + S)$ will be:

$$[R]\theta = [(A_2 + a_2)\delta^2 + (A_1 + a_1)\delta + (A_0 + a_0)]\theta = (A_2 + a_2)\ddot{\theta} + (A_1 + a_1)\dot{\theta} + (A_0 + a_0)\theta = D. \quad (34)$$

It is seen from equation (34) that the *fundamental control operator* $[S_0]$ *does not change the order of differential equation when passing from a non-controlled system* (A) *to the controlled one* $(A + S)$, *but merely modifies the coefficients of the differential equation*, viz.: original A_2 (without control) becomes $(A_2 + a_2)$ with the control operating, etc. Coefficients a_2, a_1, a_0 are called the *coefficients of intensity* of the fundamental control; they can be either positive or negative; some of them may be zero. They cannot be all three zero simultaneously, because this would be equivalent to the absence of any control.

In spite of the fact, however, that the fundamental control operator $[S_0]$ does not modify the order of the original differential equation describing the non-controlled process of system (A), the analytical form of solutions of equation (34)

for the controlled process, may be quite different from those of the original differential equation (1) without control. For instance, for a sufficiently high positive value of a_1 (assuming that $a_2 = a_0 = 0$) the controlled process may be expressed by a damped aperiodic function $Be^{-\alpha t}$, whereas, the non-controlled one ($a_1 = 0$) it may be represented by a damped oscillatory function of the form $Be^{-\alpha t} \cos(\omega t + \varphi)$.

Conversely, the latter form of solution can be made undamped ($\alpha = 0$) for a sufficiently high negative value of a_1 at which $(A_1 + a_1) = 0$ in equation (34); this happens for instance, in the case of an anti-rolling control with a reversed velocity control, when the condition of a sustained undamped self-rolling can be artificially produced in this manner.

As regards factors, which depend solely on the order of the differential equation, such factors are not modified by the fundamental control, insofar as the latter does not change the order of the original differential equation (1). Thus, for example, when both equations (1) and (34) have oscillatory solutions (i.e. when the roots of their respective characteristic equations are complex), both systems, the uncontrolled (equation 1) and the controlled (equation 34), cannot have more than one single frequency in their free oscillation because a characteristic equation of the second degree cannot have more than one pair of conjugate complex roots.

Conditions are different, however, for non-fundamental control systems considered below.

12. *Non-Fundamental Control Systems.*

In this case $[S] = [S^{\cdot}] + [S_0] + [S.]$. It is apparent that the component $[S_0]$ can be left out in this study because it leads to the fundamental control already studied in section 11. It is sufficient, therefore, to take into account only the residue operators $[S^{\cdot}]$ and $[S.]$ for this analysis.

It will be shown that the effects of these operators on the controlled process are somewhat different so that two separate cases have to be considered:

1. $[S] = [S^{\cdot}]$ and accordingly $[R] = [A + S^{\cdot}]$. (35)

2. $[S] = [S.]$ and accordingly $[R] = [A + S.]$. (36)

It will be shown now that both residue operators $[S^{\cdot}]$ and $[S.]$ *raise the order of the differential equation for the controlled process.*

For operator (35) this is self-evident because it adds higher order terms e.g. $[a_4\delta^4 + a_3\delta^3]$ to the original operator $[A]$ of the second order, so that the resultant operator $[R]$ is now:

$$[R] = [A + S^{\cdot}] = [a_4\delta^4 + a_3\delta^3 + A_2\delta^2 + A_1\delta + A_0] \quad (37)$$

and the corresponding differential equation of the controlled process is:

$$[R]\theta = a_4\theta^{IV} + a_3\ddot{\theta} + A_2\ddot{\theta} + A_1\dot{\theta} + A_0\theta = D, \quad (37a)$$

whereas the differential equation of (A) alone was:

$$A_2\ddot{\theta} + A_1\dot{\theta} + A_0\theta = D. \quad (1)$$

The controlled process $(A + S)$ will thus have new features which the original non-controlled process does not have, inasmuch as the order of corresponding differential equations is different in both cases. The control introduced by operator $[S^{\cdot}]$ is, thus, a non-fundamental one because it changes the order of the original differential equation of system (A).

In the second case when the residue operator is of the type $[S.]$ conditions are similar i.e., the control is not a fundamental one. In fact the residue operator $[S.]$ contains negative powers of δ; let $\delta^{-\lambda}$ be the lowest order term[6] in this operator. As was explained in section 8, negative powers of δ must be eliminated from the operator by multiplying the whole differential equation $[A + S.]\theta = D$ by δ^{λ}; this gives $\delta^{\lambda}[A + S.]\theta = \delta^{\lambda}D$. Since

$$[S.] = [a_{-1}\delta^{-1} + a_{-2}\delta^{-2} + \cdots a_{-\lambda}\delta^{-\lambda}]$$

one has

$$\delta^{\lambda}[A + S.] = [a_{-1}\delta^{\lambda-1} + a_{-2}\delta^{\lambda-2} + \cdots + a_{-\lambda} + A_2\delta^{\lambda+2} + A_1\delta^{\lambda+1} + A_0\delta^{\lambda}]$$

[6] Expression "lowest order term" used here corresponds to the highest order of integration inasmuch as the latter in operational notations used here, is designated by terms $a_{-4}\delta^{-4}$ with negative exponents.

and, on regrouping the terms, the corresponding differential equation of the controlled process will be:

$$A_2\theta^{(\lambda+2)} + A_1\theta^{(\lambda+1)} + A_0\theta^{(\lambda)} + a_{-1}\theta^{(\lambda-1)} + \cdots a_{-\lambda+1}\dot{\theta} + a_{-\lambda}\theta = \frac{d^\lambda D}{dt^\lambda}. \quad (38)$$

If, for instance, $\lambda = 1$, $[S.] = [a_{-1}\delta^{-1}]$, one then has:

$$\delta[A + a_{-1}\delta^{-1}]\theta = [A_2\delta^3 + A_1\delta^2 + A_0\delta + \delta\delta^{-1}a_{-1}]\theta = [\delta D]$$

and the differential equation for $(A + S.)$ will be:

$$A_2\dddot{\theta} + A_1\ddot{\theta} + A_0\dot{\theta} + a_{-1}\theta = \frac{d}{dt}D. \quad (39)$$

It is seen that the effect of the residue control operator $[S.]$ results again in *non fundamental* features in the controlled process insofar as it ultimately raises the order of the original differential equation for (A).

The difference between the control features of residue operators $[S^{\cdot}]$ and $[S.]$ is in that the higher order terms appearing in the differential equation in this manner are due to those of the *control operator itself for the former and to the terms of the* $[A]$ *for the latter.*

Furthermore, operator $[S^{\cdot}]$ does not change the form of the external disturbance D as far as its action on the controlled process is concerned, whereas the operator $[S.]$ modifies also this right-hand term which becomes now $\frac{dD}{dt}$, $\frac{d^2D}{dt^2}$ etc., according to the value of λ in the lowest order term of $\delta^{-\lambda}$ entering into $[S.]$. In any case, for both residue operators $[S^{\cdot}]$ and $[S.]$ the control is non-fundamental insofar as the order of the differential equation of the control process is raised.

Physical significance of residue operators $[S^{\cdot}]$ and $[S.]$ in connection with control problems is explained in examples 4 and 7 respectively of section 17 below.

13. *Principle of Superposition in Linear Systems: Control Terms.*

Control operators were considered so far by their general features. We shall now analyze the effect of individual terms —*the control terms*, in these operators. Inasmuch as we are concerned here only with linear operators, the effect of the whole operator $[S]$ on a linear system (A) is equal to the sum of individual effects of control terms $[S_i]$ in view of the principle of superposition applicable to linear systems. In order to simplify notations somewhat, the square brackets indicating operators will be dropped from now on, whenever no confusion with ordinary algebraic notations can occur. Furthermore, we can systematize the preceding notations by attaching to the coefficients of different operators the same lower indices which correspond to different powers of δ in the operator. We shall thus have: for the most general form of the control operator S the expression:

$$S = \sum_i S_i, \tag{40}$$

in which S_i are different control terms of the form $S_i = a_i\delta^i$. One needs to remember only that these notations are operational notations i.e., $S_i = a_i\delta^i$ in reality is $[S_i] = [a_i\delta^i]$ etc.

In order to acquire a still broader viewpoint on the whole multiplicity of possible controls, we have to consider also indirect methods of control in which the control is exerted either on derivatives $\frac{dS}{dt}, \frac{d^2S}{dt^2} \cdots$ or integrated values $\int S dt$, $\int dt \int S dt$ of the control action S, rather than on this control action S itself.

In the above notations, this can easily be taken care of by attaching upper indices to the control terms. With this system of double indices, one has, for example $S_n^0 = a_n^0\delta^n$ for the general term of a direct control operator; similarly $S_n^1 = a_n^1\delta^n$ for an indirect operator $[S^1] = \left[\frac{dS}{dt}\right]$; similarly $S_n^{-1} = a_n^{-1}\delta^n$ for an indirect operator $[S^{-1}] = [\int S dt]$ etc. The whole multiplicity of all possible controls is then contained in the

following expressions:

$$\left.\begin{aligned}
\frac{d^2S}{dt^2} = S^2 &= \cdots + a_m{}^2\delta^m + \cdots a_2{}^2\delta^2 + \\
&\quad + a_1{}^2\delta + a_0{}^2\delta^0 + a_{-1}{}^2\delta^{-1} + \cdots \\
\frac{dS}{dt} = S^1 &= \cdots + a_m{}^1\delta^m + \cdots a_2{}^1\delta^2 + \\
&\quad + a_1{}^1\delta + a_0{}^1\delta^0 + a_{-1}{}^1\delta^{-1} + \cdots \\
S = S^0 &= \cdots + a_m{}^0\delta^m + \cdots a_2{}^0\delta^2 \\
&\quad + a_1{}^0\delta + a_0{}^0\delta^0 + a_{-1}{}^0\delta^{-1} + \cdots \\
\int Sdt = S^{-1} &= \cdots + a_m{}^{-1}\delta^m + \cdots a_2{}^{-1}\delta^2 \\
&\quad + a_1{}^{-1}\delta + a_0{}^{-1}\delta^0 + a_{-1}{}^{-1}\delta^{-1} + \cdots \\
\int dt \int Sdt = S^{-2} &= \cdots + a_m{}^{-2}\delta^m + \cdots a_2{}^{-2}\delta^2 \\
&\quad + a_1{}^{-2}\delta + a_0{}^{-2}\delta^0 + a_{-1}{}^{-2}\delta^{-1} + \cdots
\end{aligned}\right\} \quad (41)$$

14. *Control Matrix and its Properties.*

We can now fix our attention on coefficients $a_j{}^k$ of these expressions by writing a two dimensional array, or matrix, of these coefficients arranged in columns according to powers of δ and in lines according to the order of S^k. Line S^0 corresponds to terms of the direct control operator and symbol $\delta^0 = 1$ is added for the sake of symmetry of notations.

$$\begin{array}{c|ccccc|c}
L \searrow & \delta^2 & \delta^1 & \delta^0 & \delta^{-1} & \delta^{-2} & \swarrow M \\
\hline
S^2 & \cdots a_2{}^2 & a_1{}^2 & a_0{}^2 & a_{-1}{}^2 & a_{-2}{}^2 & \cdots \\
S^1 & \cdots a_2{}^1 & a_1{}^1 & a_0{}^1 & a_{-1}{}^1 & a_{-2}{}^1 & \cdots \\
S^0 & \cdots a_2{}^0 & a_1{}^0 & a_0{}^0 & a_{-1}{}^0 & a_{-2}{}^0 & \cdots \\
S^{-1} & \cdots a_2{}^{-1} & a_1{}^{-1} & a_0{}^{-1} & a_{-1}{}^{-1} & a_{-2}{}^{-1} & \cdots \\
S^{-2} & \cdots a_2{}^{-2} & a_1{}^{-2} & a_0{}^{-2} & a_{-1}{}^{-2} & a_{-2}{}^{-2} & \cdots \\
\hline
\nearrow M & & & & & & \swarrow L
\end{array} \quad (42)$$

The points on all four sides of the array indicate that it can be continued indefinitely.

We have to be guided now by the physical significance of the individual control terms. It has been stated already that the operator of the controlled system is $[R] = [A + S]$; in other words, our final goal is to obtain a direct operator

$[S^0]$, or simply S^0 and not indirect operators such as $S^1 = \frac{dS}{dt}$. $S^2 = \frac{d^2S}{dt^2}$ or $S^{-1} = \int S dt$, $S^{-2} = \int dt \int S dt$.[6]

For this purpose, the line corresponding to the indirect operator S^1 *must be integrated once*, in order to obtain the corresponding *direct* operator S^0; integration means however, a multiplication by δ^{-1} and this entails shifting this line of coefficients by one order to the right because an original control term, e.g. $a_3{}^1\delta^3$ being multiplied by δ^{-1}, becomes $a_3{}^1\delta^2$, that is the coefficient $a_3{}^1$ is now in the δ^2 column and not in δ^3 column as it was originally in array (42); the same holds for all other coefficients of the line S^1. As regards S^1 itself, it becomes now a direct operator, $S_*{}^0$; the asterisk means that one constant of integration has appeared as the result of one integration; in fact $S^1 = \frac{dS}{dt}$, if one integrates, one has $\delta^{-1}S^1 = \int \frac{dS}{dt} dt = S^0 + \text{const} = S_*{}^0$ in the above notations.

The next line S^2 in a similar manner is shifted by two orders to the right. In fact, $S^2 = \frac{d^2S}{dt^2}$ by the process of double integration it becomes:

$$\delta^{-2}S^2 = \int dt \int \frac{d^2S}{dt^2} dt = S^0 + c_1 t + c_2 = S_{**}{}^0,$$

two asterisks meaning again that *two* constants of integration have appeared owing to the double integration, etc.

If one considers now the "integrated controls," e.g.

$$S^{-1} = \int S dt; \qquad S^{-2} = \int dt \int S dt,$$

these have to be differentiated, i.e. corresponding control terms situated in lines S^{-1}, $S^{-2} \cdots$ must be *multiplied by* δ, $\delta^2 \cdots$ i.e. corresponding terms of the array must be shifted

[6] It is useful to recall again that expressions S^1, S^2, S^3 etc., are operational notations, equivalent to: $[S^1] = \left[\frac{dS}{dt}\right]$, $[S^2] = \left[\frac{d^2S}{dt^2}\right] \cdots$ and *not algebraic expressions*, such as "square of S" for S^2 etc.

by one order to the left in line S^{-1}; by two orders to the left for terms in line S^{-2} etc., in order to obtain terms of the direct operator. The control operators are then transformed: S^{-1} becomes $S.^{0}$ (one point means that there was one differentiation of the whole differential equation, including D); S^{-2} becomes $S..^{0}$ etc. All control terms are now "reduced" to those of equivalent *direct* operators. Everything happens as if the original matrix (42) underwent a "shear" with respect to line S^0 which remains unchanged.

15. *Sheared Matrix of Control Terms.*

The new array of control terms of matrix (42) for the corresponding direct control operators S^0 is then:

	δ^3	δ^2	δ^1	δ^0	δ^{-1}	δ^{-2}	δ^{-3}		
S_{0***}	·	·	·	a_3^3	a_2^3	a_1^3	a_0^3	a_{-1}^3	a_{-2}^3
$S_{**}{}^{0}$	·	·	a_3^2	a_2^2	a_1^2	a_0^2	a_{-1}^2	a_{-2}^2	a_{-3}^2
$S_{*}{}^{0}$	·	a_3^1	a_2^1	a_1^1	a_0^1	a_{-1}^1	a_{-2}^1	a_{-3}^1	·
S^0	a_3^0	a_2^0	a_1^0	a_0^0	a_{-1}^0	a_{-2}^0	a_{-3}^0	·	
$S.^{0}$	a_2^{-1}	a_1^{-1}	a_0^{-1}	a_{-1}^{-1}	a_{-2}^{-1}	a^{-1}	·	·	
$S..^{0}$	a_1^{-2}	a_0^{-2}	a_{-1}^{-2}	a_{-2}^{-2}	a_{-3}^{-2}	a^{-2}	·		
$S...^{0}$	a_0^{-3}	a_{-1}^{-3}	a_{-2}^{-3}	a^{-3}	a	·	·		
A		A_2	A_1	A_0					

(43)

One can obtain now a greater degree of generalization from this "sheared" array (43) of control terms than was possible before by considering only isolated examples.

It must be noted that the resultant operator is still $[R] = [A + S]$, but S has now a greater degree of variety (S^0, $S_*{}^0$, $S_{**}{}^0 \cdots$), than before. It is useful to write also the coefficients A_2, A_1, A_0 of the operator $[A] = [A_2\delta^2 + A_1\delta + A_0]$ of system (A); these coefficients are indicated in δ^2, δ^1, δ^0 columns below the horizontal line on the lower part of the array.

The fact that the columns of array (43) are arranged according to powers of δ permits obtaining the resultant operator $[R] = [A + S]$ in a relatively simple manner.

For an operator $[A]$ of the second order, the control terms contained between vertical heavy lines of array (43)

are *fundamental control terms*, because these terms do not modify the order of the original operator $[A]$ for the "natural" uncontrolled system (A). Outside this range the control terms have properties of a non-fundamental control (section 12).

Control terms situated in the same column of array (43) correspond to the same effect in the controlled process; they can be designated for this reason as *equivalent control terms*. A more detailed study of these terms in connection with various control problems is given in the following section. In fact, in order to form $[R]$, operators $[A]$ and $[S]$ are to be added by grouping terms with the same powers of δ.

For example, consider coefficients a_1^0, a_2^1, a_3^2 in column δ^1 of array (43); the corresponding reduced control operators are $[a_1^0\delta]$, $[a_2^1\delta]$, $[a_3^2\delta]$ and the corresponding resultant operators are:

$$[A_2\delta^2 + (A_1 + a_1^0)\delta + A_0], \quad [A_2\delta^2 + (A_1 + a_2^1)\delta + A_0],$$
$$[A_2\delta^2 + (A_1 + a_3^2)\delta + A_0].$$

It is seen that the action of these control terms *is felt in the same manner* on the controlled process, namely: it modifies the damping coefficient A_1 of the original uncontrolled system; in this manner, these three particular control terms are, indeed, *equivalent*.

Equivalent terms are located on diagonals parallel to LL in the original array (42) in view of the fact that the process of "shearing" transforms the diagonals LL of array (42) into columns of array (43) and vice versa. Perpendicular diagonals MM in array (42) and, hence, the lines of array (43) do not exhibit this feature of equivalence as is easy to ascertain directly.

Similar conclusions are applicable to other columns of sheared array (43)—control terms situated in the same column are equivalent.

16. *Equivalent Control Terms. Floating Zero in a Control Process.*

The following examples in connection with the problem of ship stabilization by activated tanks illustrate the significance of equivalent control terms. From the theory of

stabilization by tanks (3), it follows that the stabilizing moment S must be controlled so as to be continuously proportional to and in phase with the angular velocity $\dot{\theta}$ of rolling, that is, it must be of the form:

$$S = \mu\dot{\theta}. \tag{44}$$

Since this moment S is a gravity moment wL of the "transferred" ballast w, L being the lever arm, clearly equation (44) is equivalent to

$$w = a\dot{\theta}, \tag{45}$$

where $a = \mu/L$.

Since the moment S (and hence a corresponding proportional quantity w) are controlled *directly* in this case, one should write instead of a, $-a_1^0$, in order to be consistent with the previous notations. The operator characterizing the control process given by equation (45) is then: $S_1^0 = [a_1^0\delta]$ in these notations. A control process of this kind can be obtained by balancing a response (e.g. in the form of a voltage of an electric circuit) obtained from an angular velocity indicating instrument, such as a constrained gyroscope (the so-called "primary" control) against a corresponding response of an instrument measuring the height of the water level in the tank, such as a float, inasmuch as this height is a measure of the amount of transfer w. This second "balancing" component in the control is usually called "secondary" or "follow up" control.

This instantaneous balance of the primary and secondary actions gives in fact a control of the form (45) or (44); in other words, it is characterized by the operator $[a_1^0\delta]$.

If one differentiates equation (45) one has:

$$\dot{w} = a\ddot{\theta}. \tag{46}$$

This is a control equivalent to the control indicated by equation (45) because the former is obtained as the result of integration of the latter in the course of time without any additional devices. According to control (46) the rate of flow $\dot{w}$ of "transferred" ballast is controlled so as to be continuously proportional to, and in phase with $\ddot{\theta}$—the angular acceleration of the ship on rolling. A control of this kind can be obtained in practice by balancing a "primary" control

action obtained from an angular accelerometer measuring $\ddot{\theta}$ against a secondary or "follow up" action derived from an instrument measuring the rate of flow $\dot{w}$ (4, 6). In order to be consistent with the previous notations, one has to attach again, corresponding indices to the a's in the operational notations. The operator characterizing the control process (46) is then:

$$S_2^1 = [a_2{}^1\delta^2]. \tag{47}$$

A multiplication by δ^{-1} i.e. integration gives:

$$\delta^{-1}[S_2{}^1] = \delta^{-1}\delta[S_2{}^0] = [S_{2*}{}^0] = \delta^{-1}[a_2{}^1\delta^2] = [a_2{}^1\delta]. \tag{48}$$

Controls given by equations (45) and (46) are, in fact, equivalent because they correspond to reduced control operators $S_1{}^0$ and $S_{2*}{}^0$ of the same form, viz.: $[a_1{}^0\delta]$ and $[a_2{}^1\delta]$ situated in the same column δ of the sheared array (43).

One can take a step further by differentiating the equation of control (46) and one has:

$$\ddot{w} = a\ddot{\theta}. \tag{49}$$

This is another form of an equivalent control. A control of this kind is obtained by balancing the response of a circuit giving $\ddot{\theta}$ against the response of a circuit measuring $\ddot{w}$. The response to $\ddot{\theta}$ is easily obtained by associating a differentiating circuit with that measuring $\dot{\theta}$. As to $\ddot{w}$, it can be obtained in different ways according to the method used for the acceleration of ballast. If the ballast is accelerated by air pressure built up in tanks, the "follow up" action can be derived from this pressure. If axial pumps are used in the cross channel, the axial thrust applied by the runner of the pump against its step bearing, can serve as a measure of $\ddot{w}$ with a sufficient degree of accuracy.

Here, again, in order to comply with the above notations, one has to attach indices to the corresponding control term, viz.:

$$S_3{}^2 = [a_3{}^2\delta^3].$$

Since this is an indirect control operator, a multiplication by δ^{-2} (i.e. double integration) is required. This gives

$$\delta^{-2}[S_3{}^2] = \delta^{-2}\delta^2[S_3{}^0] = [S_{3**}{}^0] = \delta^{-2}[a_3{}^2\delta^3] = [a_3{}^2\delta], \tag{50}$$

which shows that the reduced direct control operator $[S_{3**}{}^{0}]$ has the same form $[a_3{}^{2}\delta]$ as operators $[a_1{}^{0}\delta]$ and $[a_2{}^{1}\delta]$ corresponding to controls (46) and (48) previously studied. Two asterisks in $[S_{3**}{}^{0}]$ indicate again the appearance of two integration constants. This establishes directly the rule already found formally from the consideration of sheared array (43) namely: *coefficients* $a_1{}^{0}$, $a_2{}^{1}$ and $a_3{}^{2}$ *situated in the same column of array* (43) *characterize the equivalent controls.*[7]

This theoretical equivalence of control terms situated in the same column of array (43) is limited to a certain extent by secondary factors which we propose to analyze now. Consider the control given by equation (46) equivalent to that given by equation (45). In order to pass from the former to the latter one integration is necessary. In operational notations this is equivalent to a multiplication by δ^{-1}. The integration process introduces, however, one constant of integration indicated by one asterisk in the reduced control operator $[S_{2*}{}^{0}]$ equation (48) so obtained. Likewise for a control given by equation (49) two integrations are necessary, this means that two constants of integration appear in this manner (two asterisks in $[S_{3**}{}^{0}]$ equation (50)), and so on for other equivalent controls. The physical significance of constants of integration in this case characterizes an *uncertainty in the distribution of ballast* in the long run. This uncertainty increases if one advances in the sequence $a_1{}^{0} \rightarrow a_2{}^{1} \rightarrow a_3{}^{2}$ of equivalent controls of gradually increasing orders. In fact for the direct control given by equation (45) there is, so to say, a rigid relation between the amount of water w in a tank and angular velocity $\dot{\theta}$ of the ship, determined by the float. For indirect controls of the rate equation (46) and of the second rate equation (49) the uncertainty in the distribution of ballast in the course of time increases, because of inevitable lack of symmetry in operation of various elements of the control equipment "dead zones" etc., which may lead to an erratic shift of ballast, between the tanks in the long run. Experiment shows, however, that an equivalent control of the

[7] It is to be noted that the equivalent control terms are also those which have the same difference $(j - i)$ between the lower and upper indices in the general expression $a_i{}^{j}$ of a control term as seen from the previous examples in connection with $a_1{}^{0}$, $a_2{}^{1}$ and $a_3{}^{2}$ terms.

form (46) under certain conditions operates even better than the basic control (45); in fact it is easier in this case to control the rate of flow $\dot{w}$ than the amount w of transfer, without any serious cumulative errors owing to the inherent stability of the ship which reduces gradually these errors. This however occurs only in exceptional cases. In general, the effect of "floating" in the control process of such higher order is to be feared.

These secondary effects "of integration constants" causing "floating" of the control action S limit to a certain extent the application of equivalent control terms in practice, and necessitate a careful consideration of conditions under which such "equivalent controls" are supposed to operate in each particular case.

One interesting feature of equivalent controls is in the fact that one can couple them together, care being taken in correlating each primary element with a corresponding secondary or "follow up" element; advantages of high order controls (e.g. rapidity of response) in such a case is matched up with advantages of low order controls (definiteness of "zero" point) without any corresponding disadvantages. Compound control systems of this kind, are, however, complicated and, for this reason, seldom employed in practice, except in cases when such complications are justified by the accuracy of control so obtained.

17. *Examples of Application of Operators to the Analysis of Different Control Systems.*

Although the preceding examples give already a number of applications of the operational method in connection with equivalent terms, an additional number of examples outside the range of equivalence may be useful for illustrating a series of other features of different methods of control.

A convenient model for investigating these general features is a damped physical pendulum (system A) suitably controlled (6). Consider a pendulum A (Fig. 1) capable of oscillating about an axis O perpendicular to the plane of paper; G is the center of gravity of pendulum. On its upper part, pendulum A has a screw C capable of displacing a weight w along the x axis at right angles to the axis of oscillation by means

of a motor M suitably controlled. Instruments (not shown) responsive either to θ (angle of oscillation), or $\dot{\theta}$ (angular velocity), or $\ddot{\theta}$ (angular acceleration) of the pendulum can be attached to the pendulum and by suitable electronic circuits (not shown) can be made to control the motor M either in the number of its revolutions, which fixes the position x of

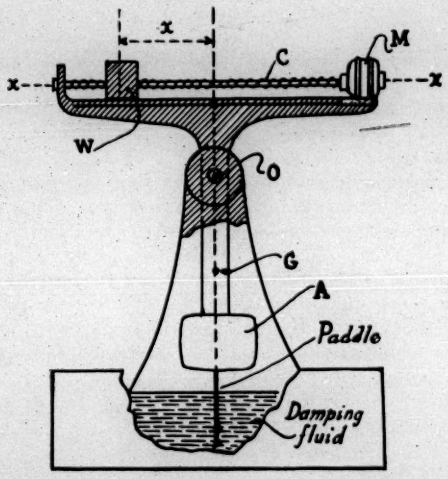


FIG. 1.

the weight w, or in speed, which establishes the speed $\dot{x}$ of weight w, or in acceleration $\ddot{x}$ etc. These different schemes can be produced experimentally for a number of combination of instruments and circuits, as explained below, omitting the details of these arrangements (4, 6). Although, strictly speaking, a system of this kind has two degrees of freedom (θ, x) which requires also a consideration of reactions between the pendulum and the weight w in a complete study of the process—for the first approximation it is sufficient to consider only the gravity moment $S = wx$, as the only "control action" (neglecting these reactions) applied to the pendulum in addi-

tion to its inherent stability moment $A_0\theta$ due to its pendulousness. This is legitimate to assume if the motion of the pendulum is relatively small and slow or when the axis x is in the same plane with the axis of oscillation; this simplification will be assumed in the following.

Operator $[A]$ of the system (A), the pendulum, in such a case has the form given by equation (2); the control operator $[S]$ will have the form $w[x]$ since w is constant. By controlling either x (lever arm), or $\dot{x}$ (speed of motor M), or $\ddot{x}$ (acceleration) in proportion to either θ, or $\dot{\theta}$, or $\ddot{\theta}$ a great variety of controlled motions can be obtained; the analysis of a few of these different motions is given below.

A universal scheme of this kind can serve as a model for a great variety of different problems of control, either dynamical (e.g. ship stabilization, automatic steering of ships, aircraft, etc.), or others; in each particular problem, one has to ascertain what values of parameters A_2, A_1, A_0 of the pendulum should be used. A few examples given below correspond to control problems frequently encountered in practical applications. The reader can easily extend this analysis to other types of controls.

1. $x = \mu\dot{\theta}$ or, which is the same, $S = a\dot{\theta}$ (μ and a are suitable coefficients of intensity of control). In this case, lever arm x is controlled in proportion to $\dot{\theta}$ which can be measured by a constrained gyroscope mounted on the pendulum. The primary control in the control circuit in this case is taken from $\dot{\theta}$ and the secondary "follow up" control is associated with the excursion x of the weight.

Ship stabilization by means of a solid moving weight suggested by Sir J. I. Thornycroft,[8] the problems of automatic

[8] It may be interesting to mention here that during his experiments on board the "Cecile" (7), Thornycroft had a control of the form: $S = a_0\theta$ (example 2 below) and not $S = a_1\dot{\theta}$. It so happened, however, that, owing to an excessive time lag in his control system, Thornycroft was able to obtain some results (about 50 per cent. of quenching). These results were due, however, not to the control $S = a_0\theta$ which he had in mind, but to a "quadrature effect" component $S_1 = a_1\dot{\theta}$ due to time lag (see Part III below).

During the experiments with a controlled pendulum at the University of Pennsylvania (1930), the writer was able to obtain a more accurate stabilization (about 90 per cent.) by introducing in the Thornycroft scheme of moving weight, a control in the form: $S = a_1\dot{\theta}$.

steering and stabilization by fins fall also in this category, if one considers the rudder angle or angle of fins, instead of the variable x considered here.

The control operator in this case is a direct operator:

$$S_1^0 = a_1^0\theta.$$

The resultant operator:

$$[R] = [A + S_1^0] = [A_2\delta^2 + (A_1 + a_1^0)\delta + A_0]$$

and the differential equation of controlled motion is:

$$A_2\ddot{\theta} + (A_1 + a_1^0)\dot{\theta} + A_0\theta = 0. \qquad (51\text{-}1)$$

Such control either increases ($a_1^0 > 0$) or decreases ($a_1^0 < 0$) the natural damping A_1 of the pendulum. If $a_1^0 < 0$ and $(A_1 + a_1^0) < 0$ the pendulum has a "negative damping," i.e. may start oscillating from zero with gradually increasing swings being given a slight initial impulse. For $(A_1 + a_1^0) > 0$ (i.e. positive damping) the weight w moves always "down hill" during the motion of pendulum; for $(A_1 + a_1^0) < 0$ (i.e. "negative" damping) weight w moves "up hill," which thus continuously increases the potential energy of the system. This is done at the expense of electrical energy absorbed by the motor; the energy of "self oscillation" in such a case is supplied by the motor and the control of the form $S_1^0 = a_1^0\dot{\theta}$ merely times it so that this transformation of energy (electrical energy → potential energy → kinetic energy of self oscillation) becomes possible.

2. Control of the form $S = a\theta$ is obtained by using the angle θ of the pendulum as primary control and the excursion of weight x as "follow up" control.

The control operator in this case is $S_0^0 = a_0^0\theta$ and the resultant operator:

$$[R] = [A + S_0^0] = [A_2\delta^2 + A_1\delta + (A_0 + a_0^0)].$$

The corresponding differential equation is:

$$A_2\ddot{\theta} + A_1\dot{\theta} + (A_0 + a_0^0)\theta = 0. \qquad (51\text{-}2)$$

It is seen that the action of control in this case merely modifies the "pendulousness" A_0 of the pendulum, either increasing it ($a_0^0 > 0$), or decreasing it ($a_0^0 < 0$) according to the connec-

tion of the control circuit. This particular control does not modify the damping,[9] but only changes the period of the pendulum.

3. Control of the form $S = a\ddot{\theta}$ is obtained by using the response of an angular accelerometer, mounted on the pendulum, as primary control and by balancing it by a follow up control, derived from the excursion x of the weight.

In this case $S_2^0 = a_2^0\ddot{\theta}$ and

$$[R] = [(A_2 + a_2^0)\delta^2 + A_1\delta + A_0].$$

The corresponding differential equation in this case is:

$$(A_2 + a_2^0)\ddot{\theta} + A_1\dot{\theta} + A_0\theta = 0. \qquad (51\text{–}3)$$

Everything happens in this case as if the moment of inertia A_2 of pendulum were modified by such a control. A control of this kind changes both the period of oscillation of the pendulum and its exponential decrement (damping). A control of this kind is of small practical interest because of the complicated form of its action above referred to.

These three examples cover the range of fundamental controls.

4. Control of the form $S = a_4\theta^{IV} + a_3\dddot{\theta}$.

A control of this kind can be obtained by using a combination of two subsequent differentiating circuits associated with angular accelerometer ($\ddot{\theta}$) as primary control and by using a follow up control from the excursion x of the weight w as in the first three cases. The control operator in this case is

$$[S.] = [a_4^0\delta^4 + a_3^0\delta^3] \text{ and}$$
$$[R] = [a_4^0\delta^4 + a_3^0\delta^3 + A_2\delta^2 + A_1\delta + A_0] = [S'] + [S_0],$$

using the notations of section 12. This control is, thus, a non-fundamental one, because the order of $[R]$ is now higher than that of $[A]$. This particular operator is precisely the residue operator $[S']$ defined in section 12. The differential equation of the controlled process is then:

$$a_4^0\theta^{IV} + a_3^0\dddot{\theta} + A_2\ddot{\theta} + A_1\dot{\theta} + A_0\theta = 0. \qquad (51\text{–}4)$$

This differential being now of the fourth order, its characteristic equation has now four roots and the additional pair of

[9] See footnote, page 481, in this connection.

roots, for suitable values of coefficients a_4^0 and a_3^0 may be a pair of conjugate complex roots; for this particular condition, the pendulum acquires an additional frequency. As the result of this, the pendulum will oscillate in a rather complicated manner, due to the superposition of two different frequencies in its free oscillation. A practical application of this particular control is indicated in section 22 below.

5. Control of the form $S_1^1 = a_1^1\theta$. This control can be obtained by establishing a "follow up" relation between the *speed* of motor M and the response of an angular velocity responsive instrument. The control operator S_1^1 in this case is an indirect one. Multiplication by δ^{-1} converts it into a direct operator. We have:

$$\delta^{-1}[S_1^1] = \delta^{-1}\delta[S_1^0] = [S_{1*}^0] = [a_1^1]$$

and $[R] = [A_2\delta^2 + A_1\delta + (A_0 + a_1^1)]$ whence the differential equation:

$$A_2\ddot{\theta} + A_1\dot{\theta} + (A_0 + a_1^1)\theta = 0. \qquad (51\text{–}5)$$

It is seen that $[R]$ in this case has the same form as in case 2. In other words, a control in which the *rate* of control action is controlled in proportion to θ has the same effect on the pendulum (i.e. is equivalent), as a control in which this action S is controlled directly in proportion to θ, which is, more or less, obvious.

The only difference between the two cases, 2 and 5, is in that in case 5, owing to integration (δ^{-1}) there will be a certain "floating" observed due to "integration constant" (asterisk) characterizing a cumulative error inherent in the "rate" control.

6. Control of the form $S_2^1 = a_2^1\ddot{\theta}$. A control in this case is the same as in case 5 only an angular accelerometer ($\ddot{\theta}$) is substituted for the constrained gyro ($\dot{\theta}$).

Corresponding operator $S_2^1 = [a_2^1\delta^2]$ is to be reduced to a direct operator. This gives:

$$\delta^{-1}[S_2^1] = \delta^{-1}\delta[S_2^0] = [S_{2*}^0] = [a_2^1\delta]$$

and $[R] = [A_2\delta^2 + (A_1 + a_2^1)\delta + A_0]$ whence the differential equation:

$$A_2\ddot{\theta} + (A_1 + a_2^1)\dot{\theta} + A_0\theta = 0. \qquad (51\text{–}6)$$

This is a control equivalent to the case 1, with the additional feature of "floating" due to "integration constant" (asterisk), unless there exists an inherent stability in system (A) sufficient to wipe out this undesirable effect. A control of this kind has been described in section (16), equation (46), in connection with the problem of anti-rolling stabilization of ships (4, 6).

7. Control of the form $S_0^1 = a_0^1\theta$. This control can be obtained by associating a response from the angle θ of the pendulum used as primary control with a follow up response proportional to the speed of motor M.

Control operator $S_0^1 = [a_0^1]$ being an indirect one must be converted into a direct one by multiplying it by δ^{-1}. This gives:

$$\delta^{-1}[S_0^1] = \delta^{-1}\delta[S_0^0] = [S_{0*}^0] = [a_0^1\delta^{-1}].$$

This is a non-fundamental "residue operator" of the type $[S.]$ (section 10). We have:

$$[R] = [A_2\delta^2 + A_1\delta + A_0 + a_0^1\delta^{-1}].$$

In this form, operator $[R]$ contains a negative power of δ, so that a multiplication of the whole differential equation, including the right-hand term D, by δ (i.e. differentiation) is necessary, in order to eliminate δ^{-1} in the operator (section 8). In the first six examples, it was assumed that $D = 0$. In fact, in these examples, the case when $D \neq 0$ could easily be taken into account by adding a corresponding particular solution of the differential equation; in this case, involving a negative power of δ, the interesting feature of control lies precisely in the case when $D \neq 0$. The differential equation of the controlled process is then:

$$\delta[R]\theta = \delta D$$

or, written explicitely, it is:

$$A_2\ddot{\theta} + A_1\ddot{\theta} + A_0\dot{\theta} + a_0^1\theta = \frac{dD}{dt}. \qquad (51\text{-}7)$$

From this example, one can see the physical significance of a non-fundamental residue operator of the type $[S.]$ investigated at the end of section 12.

Assume that D = Const.; in such a case $\frac{dD}{dt} = 0$; equation (51–7) indicates that the behavior of a controlled system in this case is the same as in the case when $D = 0$. In other words, the controlled system in this case will behave in the same manner whether we add to the pendulum a certain weight w_0 at a constant distance x_0, or not. In fact, a particular solution of equation (51–7) will be zero in both cases.

From a purely physical consideration of this case, the above mentioned condition is possible only in the case when the controlled weight w will automatically go to such position x_1 at which $wx_1 = w_0x_0$ whereby the effect of a constant moment is neutralized. This is exactly what happens in this case. In fact the angle θ of the pendulum in this case controls *the rate* at which the controlled weight w is moving. Hence as long as the pendulum is not on the vertical, this weight w is moving. Conversely the weight w ceases to move only when the pendulum becomes vertical. The automatic compensation of a constant disturbing moment w_0x_0 by the controlled moment wx is possible here, because in this case *there exits no fixed relation between the position of pendulum and position of the weight w.*

An interesting application of a control of this kind, involving negative power δ^{-1} in the resultant operator, was made a few years ago in Germany in connection with the automatic steering problem.[10] In commonly encountered automatic steering devices the directional control is generally connected to a gyroscopic compass so that a definite relation between rudder angle ρ and the deviation θ from the set course is obtained. In addition to this there are also derivatives $\dot{\theta}$, $\ddot{\theta}$ present in the control, but these are of no interest for this particular study, so that we can assume that the control action S is of the form:

$$S_0^0 = \kappa\rho = a_0^0\theta. \qquad (52)$$

If a constant disturbance, such as wind on the beam for

[10] A theoretical analysis of this condition was made (8) however a few years prior to the application of this particular method of control to the automatic steering problem.

example, acts on the ship causing a steady moment D_0 to be applied to the ship in azimuth, the "weather helm" ρ_0 checking this disturbance (so that $\kappa\rho_0 = D_0$) can be made only *at the expense of a deviation* θ_0 *from the set course.* In other words in order to be able to counteract D_0 the ship will have to fall off *her predetermined course in view of the rigid connection between* rudder and the course established by a control of this type. With a control of rudder according to operator $[S_0^1] = [a_0^1]$, leading to the appearance of a term $[a_0^1\delta^{-1}]$ in the resultant operator $[R]$ this rigid relation between the rudder angle and the departure from the set course is eliminated as was shown in connection with the pendulum. In fact the rudder in this case will continue to move so that finally the "weather helm" will be introduced automatically so as to compensate for D_0 while the ship is following a correct course ($\theta = 0$).

These few examples in connection with pendulum control give a sufficient illustration to different control problems encountered in practice. Investigation of other cases follows the same general procedure, viz.:

(1) One forms the expression of the control operator $[S]$.
(2) One forms the expression of the resultant operator $[R]$.

If $[R]$ does not contain negative powers of δ, the differential equation of the controlled process in then $[R]\theta = D$.

If $[R]$ does contain such negative powers, say, of the lowest order λ the differential equation of the controlled process is: $\delta^\lambda[R]\theta = \delta^\lambda D$ in operational notations.

(*Continued in December Issue.*)

A New Quick Freezing System.—LUIS H. BARTLETT AND H. E. BROWN. (*Refrigerating Engineering* Vol. 42, No. 2.) Freezing foodstuffs by direct contact with a chilled heat transfer fluid has several outstanding advantages over other methods. In the University of Texas "polyphase" quick freezing system the heat transfer medium employed is a chilled solution of a freezing point depressing solute, through which is dispersed the solid phase of the solute in finely subdivided form. One example of this type of medium is an invert sugar solution which has been agitated and simultaneously chilled until a disperse solid phase of very small ice crystals has been formed. Heat transfer from foodstuffs to this medium is very fast and is nearly independent of fluid velocity. The mechanical element of the system comprises a refrigerated surface in contact with the medium, an agitator which operates close to this surface, and suitable means to secure contact between the chilled medium and the foodstuff and also to effect separation when freezing has been accomplished. Probably the simplest form of equipment is a refrigerant jacketed horizontal tube, which is partly filled with medium and inside which a screw conveyor operates. This is also equipped with hoppers at each end, a return tube to permit continuous circulation of the medium, and a screen or grid to effect separation of the frozen food from the stream of medium. In operation, food is dropped upon the surface of the polyphase medium in the feed hopper, where it floats partially submerged. From there it is carried into the screw conveyor by the motion of the stream of medium. This stream serves as a perfectly elastic conveyor which cannot deform or crush the most delicate portions of food. The speed of the conveyor may be regulated to allow complete freezing of various sized articles during their travel through the tube. An unusual drive gear, which rotates the screw about one-fifth turn forward and then about one-sixth turn backward, is employed. After essentially complete freezing has occurred and the individual pieces have acquired considerable mechanical strength, the stream of medium flows through the screen and returns to the feed hopper. This screen is raked periodically to remove the frozen product, which is then drained and packaged.

R. H. O.

NOTES FROM THE NATIONAL BUREAU OF STANDARDS.*

ISOLATION OF HYDROCARBONS IN KEROSENE FRACTION OF PETROLEUM.

A second report of the work on hydrocarbons in the kerosene fraction of petroleum, which is being carried on at the Bureau as part of the American Petroleum Institute Research Project 6, appears as RP1423 by Beveridge J. Mair and Anton J. Streiff, in the October number of the *Journal of Research.*

Four hydrocarbons, 1,2,3,4-tetramethylbenzene, 5,6,7,8-tetrahydronaphthalene, 2-methyl-5,6,7,8-tetrahydronaphthalene, and 1-methyl-5,6,7,8-tetrahydronaphthalene, were separated by azeotropic distillation, followed by crystallization, of narrow-boiling fractions of aromatic material which (*J. Research N. B. S.*, **24**, 392 (1940) RP1289) had been separated by extraction and absorption from that part of the petroleum which distilled between 114° to 144° C. at 56 mm. Hg (corresponding approximately to 200° to 230° C. at 760 mm. Hg). "Best lots of each of these four hydrocarbons from petroleum were prepared, and synthetic lots of three of them were purified for purposes of identification and determination of properties. The amounts of impurity in these various lots of material were determined as follows, in mole fraction:

Hydrocarbon.	From Petroleum.	From Synthetic Material.
1,2,3,4-Tetramethylbenzene	0.0006 ± 0.0001	—
5,6,7,8-Tetrahydronaphthalene	0.015 ± 0.002	0.0057 ± 0.0008
1-Methyl-5,6,7,8-tetrahydronaphthalene	0.0029 ± 0.0004	0.0034 ± 0.0005
2-Methyl-5,6,7,8-tetrahydronaphthalene	0.0046 ± 0.0007	0.031 ± 0.004

Values of certain physical properties of these four hydrocarbons, all extrapolated to zero impurity, were determined

* Communicated by the Director.

as follows:

Property.	1,2,3,4-Tetramethylbenzene.	5,6,7,8-Tetrahydronaphthalene.	1-Methyl-5,6,7,8-tetrahydronaphthalene.	2-Methyl-5,6,7,8-tetrahydronaphthalene.
Boiling point at 760 mm., ° C.	205.04 ± 0.03	207.57 ± 0.10	234.35 ± 0.05	220.03 ± 0.05
Freezing point in air at 1 atm., ° C.	−6.25 ± 0.01	−35.80 ± 0.02	−22.90 ± 0.03	−30.75 ± 0.03
Density at 25° C., g./ml.	0.9015 ± 0.0001	0.9662 ± 0.0002	0.9683 ± 0.0002	0.9500 ± 0.0002
Refractive index at 25° C.				
n_C	1.51362 ± 0.00007	1.53466 ± 0.00015	1.53738 ± 0.00010	1.52920 ± 0.00010
n_D	1.51811 ± 0.00007	1.53919 ± 0.00015	1.54190 ± 0.00010	1.53365 ± 0.00010
n_F	1.52923 ± 0.00007	1.55065 ± 0.00015	1.55325 ± 0.00010	1.54498 ± 0.00010
Specific dispersion at 25° C.				
$(n_F - n_C)/d$	0.01732 ± 0.00005	0.01655 ± 0.00010	0.01639 ± 0.00005	0.01662 ± 0.00005

MILDEW RESISTANCE OF COTTON BAGS.

The decision of the War Department to use cotton instead of jute for sand bags has raised several new questions for consideration, one of which is the ability of the cotton to resist attack by mildew. Mildew spores are a common constituent of the atmosphere, and readily attack cotton when the moisture conditions are right. Preventatives, such as copper naphthanate, have been shown to be satisfactory and are in commercial use, and a requirement for mildew-proofness is included in a few Federal specifications. Formerly, tests for this property were made for the Bureau by the Department of Agriculture where many years have been spent in developing mildew-proofing agents and testing methods. However, the job of testing sand bags for the Army is so big, and the demand for speed so great, that the former method of operation is impractical. The Bureau can no longer make all of the other tests and farm out the mildew test to the Department of Agriculture. Therefore, an agreement has been reached with that department under which they will continue their research and development work; but the Bureau will do all of the testing. This has involved setting up a new laboratory where sterile conditions can be maintained, and providing additional equipment and personnel.

EFFECTS OF SALTS ON QUINONE SULFONATES.

Certain dyes, called indicators, change color markedly when their ionization is affected by the presence of acids or bases. Chemists, therefore, use them in estimating acidity or hydrogen ion concentration. However, even neutral salts effect a small change in the color of dyes, independent of their ionization, and strong acids behave like neutral salts in producing this secondary effect. This must be taken into account in precise determinations of hydrogen ion concentration.

An investigation by Baker Wingfield and S. F. Acree is concerned with these secondary effects of acids and salts. As explained in the report of this work (RP1424) in the October *Journal of Research*, it was initiated on an organic compound, β-naphthoquinonsulfonic acid, which, though not an indicator, contains certain groups common to a large class of indicators, but is less complex than the useful members of that group. Thus, overlapping absorption bands were avoided, and the measurements of the salt effects were simplified.

The absorption of light by the dye increased with increase of concentration of the salt or acid. In the presence of acids and salts, new absorption bands for the dye did not develop in the visible or ultraviolet. The data suggest that the specific effects of ions on the absorption indices of positive, neutral, and negative forms of an indicator must be taken into account, as well as the normal effects of the ions on the dissociation of the indicator, for accurate pH measurements.

SOIL CORROSION STUDIES.

Kirk H. Logan, chief of the Bureau's Soil Corrosion Section, is now on his biennial trip for the removal of soil-corrosion specimens and burial of new ones. The samples which are being placed in the ground this fall include six varieties of cast iron containing different percentages of nickel, copper, and carbon; three alloy steels; and a special lead alloy. At some test sites, sets of zinc-iron couples are also being installed, on which observations will be made to determine the

current densities required for protecting pipes by the use of zinc.

Interest in this subject is growing, and the application of the results of the soil-corrosion studies has been widespread. By invitation, papers on different phases of the work were presented at a special research conference on corrosion which was held during August at Gibson Island, Md., under the auspices of the National Research Council. The Federal Power Commission is particularly concerned with these studies because of their bearing on the valuation of pipe lines used for transmitting gas. One of the Commission's engineers spent over two weeks at the Bureau making measurements on samples of soil.

THE FRANKLIN INSTITUTE.

STATED MONTHLY MEETING, 8:15 P.M., WEDNESDAY, OCTOBER 15, 1941.

The stated monthly meeting for October of The Franklin Institute was held in the Hall of the Institute at 8:15 P.M. on the evening of Wednesday, October fifteenth.

Mr. Walton Forstall, Vice-President, presided.

The meeting was called to order and announcement made that minutes of the last meeting (Medal Day) were published in full in the September JOURNAL. These were approved as printed.

The Chair then called upon the Secretary of the Institute, Dr. Henry Butler Allen, for his report.

Dr. Allen stated that since the last report on membership the following additions had been made:

Sustaining	1
Active	38
Associate	35
Student	103
Total	177

With the approval of the Board of Managers, the Secretary then submitted the name of George Simpson Eddy for Honorary Membership in The Franklin Institute, "in consideration of many years of devoted research into the life and writings of Benjamin Franklin, pursued with the scientific spirit of truth, resulting in a unique and invaluable fund of information interpreting Franklin's greatness."

Upon motion, duly seconded and carried, Mr. Eddy was elected an Honorary Member.

Mr. Forstall then announced that the speaker of the evening was Charles Camsell, C.M.G., LL.D., F.R.S.C., Deputy Minister of Mines and Resources, Canada, and Commissioner of the Northwest Territories. Dr. Camsell, noted mining engineer and explorer, was awarded, among many other honors, the Murchison Grant by the Royal Geographical Society for exploration, in 1922; and the Gold Medal of the Institution of Mining and Metallurgy in 1931.

The speaker announced as his subject "The Discovery and Development of Radium Deposits of Great Bear Lake, N.W.T.," and gave a most interesting account of Canada's mining history, dwelling particularly upon the discovery and development of radium-bearing deposits at Great Bear Lake, situated on the fringe of the Arctic Circle in Northern Canada. He described the problems of mining ores in an arctic climate; the hardships of livelihood there; the difficulties of transportation, and the final triumph of engineers over all these obstacles.

At the end of his talk Dr. Camsell showed an interesting motion picture descriptive of the life at Great Bear Lake, mining the radium deposits, and finally its transportation 1500 miles by water and 2500 miles by rail to the reduction works on Lake Ontario.

The meeting was adjourned by a rising vote of thanks to the speaker for a very enjoyable and instructive evening.

HENRY BUTLER ALLEN,
Secretary.

LIBRARY NOTES.

Photostat Service. Photostat prints of any material in the collections can be supplied on request. Orders received in the morning are filled the same day. The average cost for a print 9 × 14 inches is thirty-five cents.

The library and reading room are open on Mondays, Tuesdays, Fridays and Saturdays from nine o'clock A.M. until five o'clock P.M., Wednesdays and Thursdays from two until ten o'clock P.M.

RECENT ADDITIONS.

AERONAUTICS.

Aerosphere—1941.

JORDANOFF, ASSEN. Safety in Flight. 1941.

JUDGE, A. W. Aircraft Engines. Volume Two. 1941.

MILLIKAN, CLARK B. Aerodynamics of the Airplane. 1941.

AGRICULTURE.

KELLOGG, CHARLES E. The Soils that Support Us. 1941.

ARCHITECTURE.

RAMSEY, CHARLES GEORGE, AND HAROLD REEVE SLEEPER. Architectural Graphic Standards. Third Edition. 1941.

BIOGRAPHY.

FAGAN, GEORGE V. Alexander Dallas Bache, Educator. 1941.

CHEMISTRY AND CHEMICAL TECHNOLOGY.

BENNETT, H., Editor-in-Chief. The Chemical Formulary. Volume 5. 1941.

BERKMAN, SOPHIA, AND GUSTAV EGLOFF. Emulsions and Foams. 1941.

Chemical Rubber Publishing Co. Handbook of Chemistry and Physics. Twenty-fifth Edition. 1941.

DENAVARRE, M. G. Chemistry and Manufacture of Cosmetics. 1941.

FIESER, LOUIS F. Experiments in Organic Chemistry. Second Edition. 1941.

KOLTHOFF, I. M., AND H. A. LAITINEN. pH and Electro Titrations. Second Enlarged and Revised Edition. 1941.

MARTIN, RAY C. Lacquer and Synthetic Enamel Finishes. 1940.

MATTIELLO, JOSEPH J., Editor. Protective and Decorative Coatings. Volume 1. 1941.

Organic Syntheses. Volume 21. 1941.

RAY, FRANCIS EARL. Organic Chemistry. 1941.

SHERMAN, HENRY C. Chemistry of Food and Nutrition. Sixth Edition. 1941.
TWYMAN, F. The Spectrochemical Analysis of Metals and Alloys. 1941.
WEISER, HARRY BOYER. Colloid Chemistry. 1939.
WENNER, RALPH R. Thermochemical Calculations. First Edition. 1941.
WOODMAN, A. G. Food Analysis. Fourth Edition. 1941.

ELECTRICITY AND ELECTRICAL ENGINEERING.

DOLE, MALCOLM. The Glass Electrode. 1941.
Electrochemical Society. Transactions. Volume 79. 1941.

ENCYCLOPEDIAS.

Britannica Book of the Year. 1941.

ENGINEERING.

HOLMSTROM, J. EDWIN. Records and Research in Engineering and Industrial Science. 1940.
JACOBY, HENRY S., AND ROLAND P. DAVIS. Foundations of Bridges and Buildings. Third Edition. 1941.
JONES, FRANKLIN D. Engineering Encyclopedia. Two Volumes. 1941.

GEOLOGY.

RICE, C. M. Dictionary of Geologic Terms. 1940.
STUTZER, OTTO. Geology of Coal. Translated and Revised by Adolph C. Noé. 1940.

MANUFACTURES.

CHAMBERLAIN, PAUL M. It's About Time. 1941.
HESS, KATHARINE PADDOCK. Textile Fibers and Their Use. Third Edition, Revised. 1941.
MCCASLIN, HERBERT J. Wood Patternmaking. 1941.
THAYER, GORDON B. Plastics Mold Designing. First Edition. 1941.

MATHEMATICS.

HODGE, W. V. D. The Theory and Applications of Harmonic Integrals. 1941.

MECHANICAL ENGINEERING.

FRENCH, THOMAS E. A Manual of Engineering Drawing. Sixth Edition, Revised and Enlarged. 1941.

METEOROLOGY.

PETTERSSEN, SVERRE. Introduction to Meteorology. First Edition. 1941.

MINING AND METALLURGY.

ALLEN, ROY M. The Microscope in Elementary Cast Iron Metallurgy. 1939.
MEIGH, C. H. The Practical Application of Aluminium Bronze. 1941.
PEELE, ROBERT, Editor. Mining Engineers' Handbook. Third Edition. Two Volumes. 1941.

NAVAL ARCHITECTURE AND NAVIGATION.

CHAPMAN, CHARLES F. Piloting, Seamanship and Small Boat Handling. 1942 Edition. 1941.
GAROCHE, PIERRE. Stowage. 1941.

PHARMACY.

SILVERMAN, MILTON. Magic in a Bottle. 1941.

PHOTOGRAPHY.

JOHNSON, ROBERT. The Art of Retouching and Improving Negatives and Prints. Fourteenth Edition, Revised and Enlarged by Arthur Hammond. 1941.

PHYSICS.

PERKINS, HARRY A. College Physics. Abridged. 1941.

ROBERTSON, JOHN KELLOCK. Introduction to Physical Optics. Third Edition. 1941.

STEWART, OSCAR M., AND BURTON L. CUSHING. Physics for Secondary Schools. Revised Edition. 1941.

SCIENTIFIC ESSAYS.

KRAUSKOPF, KONRAD BATES. Fundamentals of Physical Science. First Edition. 1941.

STEAM ENGINEERING.

JACKSON, PERCY G. Boiler Feed Water. Third Edition, Revised and Enlarged. 1935.

SUGAR.

Sugar Reference Book and Directory. 1941.

NOTES FROM THE BARTOL RESEARCH FOUNDATION.

THE PRODUCTION OF NEUTRONS BY THE COSMIC RADIATION.*

BY

S. A. KORFF.

The present state of knowledge about cosmic-ray neutrons is reviewed. The altitude dependence of the rate of production is investigated, employing neutron counters and the radio-balloon technique. A rate of increase with elevation is found which is equal to that of the soft component. Other experiments are described in which the coincidence between neutrons and showers is observed at two elevations, by observing the simultaneous discharge of neutron and shower counters. The calibrations and controls on the experiments are described. It is further shown that a cross section of less than 10^{-25} cm. sq. will suffice to account for the observed number of neutrons if they are produced by photons. In summary, the altitude dependence, the observed coincidences and the cross section considerations all point to a production of the neutrons by the soft component presumably by a nuclear photodisintegration process.

THE COSMIC-RAY LATITUDE EFFECT AND THE SINGLE PRIMARY COMPONENT HYPOTHESIS.†‡

BY

W. F. G. SWANN.

On the assumption that a single primary component produces mesotrons at a definite altitude in the stratosphere,

* An abstract of an article published in full in the *Proceedings of the American Philosophical Society*, **84**, 589 (1941).

† A portion of a paper presented at the Providence, Rhode Island meeting of the American Physical Society, June 20–21, 1941, under the title, "Consequences of the assumption of a single primary component in the cosmic radiation."

‡ Reprinted from *The Physical Review*, **60**, 470 (1941).

the writer finds that, if $F(E)dE$ represents the energy distribution of the vertically directed mesotrons at the point of production, then the number of mesotrons at a distance x below this point, and remaining from the group $F(E)dE$ at $x = 0$ is Δn_x, where

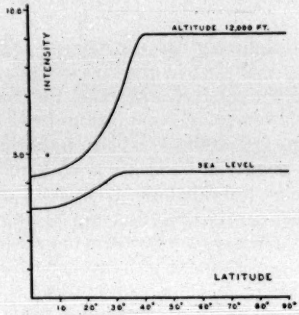
$$\Delta n_x = [(1 + A/(\gamma + 1))e^{-\lambda x} - A/(\gamma + 1)]^{[\lambda c\tau_0(\gamma+1)+A]-1} F(E)dE, \quad (1)$$

with $A \equiv \alpha\rho_0/\lambda m_0c^2$, and $\gamma \equiv E/m_0c^2$; m_0, τ_0, ρ_0, α being, respectively, the rest mass of the mesotron, its mean life at rest, the atmospheric density at $x = 0$, the ionization energy loss per unit of mass (per unit area) traversed, and where λ is the constant in the expression $\rho = \rho_0 \exp[\lambda x]$ for the atmospheric density at the depth x.

If we regard the mesotron intensity as responsible for the sea-level latitude effect, we must suppose that the minimum energy of the mesotrons at magnetic latitude 35°, where the increase of intensity with latitude ceases, must be about 10^9 ev., since that is the energy necessary to penetrate the atmosphere. If the primaries are protons, the energy necessary to penetrate the earth's magnetic field at 35° latitude is 5×10^9 ev., so that for equal distribution of energy among the mesotrons, each proton would give rise to five mesotrons. If the energy distribution law of the protons followed the inverse cube form, the numbers contained above the lower magnetic limit E_m for vertical entry would be proportional to $1/E_m^2$, so that since E_m is 10^{10} ev. for the equator, the ratio of the proton intensity at latitude 35° to that at 0° (equator) would be 4 : 1. This would also be the ratio for the corresponding mesotron intensities at latitudes 35° and 0°, for all altitudes, if mean-life considerations were neglected.

However, mesotrons of 2×10^9 ev. corresponding to the lower limit for mesotrons at the equator would have a mean life of less than 4×10^{-5} sec., averaged over the path concerned, and, if produced at an altitude of 20 kilometers, only about 0.2 of them would reach sea level. The corresponding mean life for the lower limit at latitude 35° would be smaller, and would result in a smaller fraction of the mesotrons to reach sea level there.

Equation (1) serves as a basis for submitting the above considerations to more exact calculation, and properly integrating over the total energy spectrum. Figure 1 shows the result of the assumption above cited to the effect that a proton



Latitude effect calculated with formula (1), and the values $\tau_0 = 2 \times 10^{-8}$ sec., and $m_0 = 160 \times$ mass of electron.

splits into five mesotrons at an altitude of 20 kilometers. The figure shows a 30 per cent. variation of the intensity from the minimum to the upper limit for the sea-level latitude effect, and shows an increase of latitude effect with altitude in accord with experimental findings. A slight modification of the constants, such as ionization energy loss, would bring the calculations into numerical agreement with experiment, while the intention here is merely to demonstrate the general picture.

Also, the details of the assumptions made, and in particular that of the splitting of the proton into five mesotrons of equal energy, are susceptible to wide variations; and indeed, such variations are demanded if we are to secure a proportion of mesotrons of practically zero kinetic energy at $x = 0$, and such as is necessary to account for the flatness of the intensity-zenith angle curve for high altitudes.

NOTES FROM THE BIOCHEMICAL RESEARCH FOUNDATION.

The Alcoholic Splitting of Proteins.—Dr. Saul M. Trister, Research Assistant in the Frick Chemical Laboratories of Princeton University, addressed the Seminar group of the Biochemical Research Foundation June 26 on "The Alcoholic Splitting of Proteins." In a historical summary Dr. Trister stated that Emil Fischer's work in 1901 on the hydrolysis of proteins provided the basis for further investigation into the quantitative relationship of the monoamino acids in proteins. When some ten years later various investigators applied saturated alcoholic hydrogen chloride instead of aqueous hydrochloric acid for protein scission, controversy arose as to whether the reaction was a hydrolysis or an alcoholysis. Some investigators felt that since scission was found to take place in absolute alcohol containing hydrogen chloride, alcoholysis might occur at the peptide link:

$$\begin{array}{l} R{-}CO{-}NHR_1 + EtOH \\ \qquad\downarrow \\ RCOOEt + H_2NR_1 \end{array}$$

Pribram opposed this idea and in 1911 suggested that alcoholysis was a secondary step following a primary hydrolysis which was due to the presence of a small amount of water:

$$\begin{array}{l} R{-}CO{-}NHR_1 + H_2O \rightarrow RCOOH + H_2NR_1 \\ \qquad\qquad\qquad\qquad\qquad\qquad\quad \downarrow \text{EtOH} \\ \qquad\qquad\qquad\qquad\qquad\quad RCOOEt + H_2O \end{array}$$

He proposed the hypothesis that scission may occur only in those places of the protein molecule where long hydrocarbon chain amino acids are attached through a fairly weak peptide linkage. Pfannl thought that if a difference could be demonstrated between the behavior of a protein when it was treated with aqueous hydrogen chloride and when alcoholic hydrogen

chloride was used, a new method for the study of proteins by preferential scission might be developed. However, he was not able to discover experimentally any difference in the end products using the two reagents.

Abderhalden carried out a number of experiments which appeared to substantiate Pfannl's original idea. From his experiments he concluded that polypeptides in a partially hydrolyzed protein hydrolysate would not split during esterification and so could be separated from the esters of monoamino acids by distillation. Pacsu and Wilson were able to discover a fallacy in his results, inasmuch as he evidently did not realize that he was introducing a considerable amount of water into his supposedly absolute alcohol by the reaction:

$$CH_3OH + HCl \rightarrow CH_3Cl + H_2O$$

and also that some splitting of the polypeptides must have taken place during esterification, since his amino nitrogen determinations were high.

Alcohol alone was found by Granacher in 1925 to split proteins at elevated temperatures. Ten years later Christomanos reported, as a result of studies on the action on proteins of various alcohols at different temperatures, that splitting occurred much more quickly with methyl than with ethyl alcohol, while isoamyl alcohol reacted more slowly than either of the others at low temperatures. At elevated temperatures, splitting occurred in all cases at about the same rate and to the same extent.

To answer the question whether the splitting was a hydrolysis or an alcoholysis, Glasstone and Hammel ("Notes from the Biochemical Research Foundation," JOURNAL OF THE FRANKLIN INSTITUTE, **231**, 499, 1941) have suggested recently that if the reaction were a hydrolysis, then extrapolation of its rate constant to zero water concentration would be zero. Their experiments showed that the reaction was a true alcoholysis, since the rate of splitting was ten times faster with absolute ethyl alcoholic hydrogen chloride than with aqueous hydrogen chloride, and decreased when the concentration of alcohol was reduced by the addition of water.

At Princeton in the fall of 1940 a group of workers became interested in the structure of proteins. They found that most

investigators accept one of two theories which have been proposed to account for most of the known facts of protein behavior. These are: (1) the postulate of ring systems as exemplified by the cyclol theory of Wrinch, and (2) the folded (or extended) chain theory which is best interpreted in the periodicity hypothesis of Bergmann.

Bergmann's periodicity theory asserts that amino acids are distributed throughout the entire peptide chain at regular intervals, i.e., each amino acid recurs with a constant whole number frequency. Using the equation:

$$p = \frac{100 \times \text{Mol. Wt.}}{\% \text{ Amino Acid} \times \text{Average Mol. Wt.}}$$

Bergmann calculated the periodicity of certain amino acids in gelatin hydrolysate, namely glycine (25.5 per cent.), proline (19.8 per cent.), and hydroxy-proline (14.4 per cent.), to be every third, sixth and ninth molecule, respectively, so that the formula for gelatin could be written either as:

G P X G X X G P X G X X

or

G X P G X X G X P G X X.

The latter form appears more likely in view of recent work by Grassman in isolating from gelatin a tripeptide which he showed to be lysyl prolyl glycine.

Experimentation was undertaken at Princeton, under the direction of Professor Pacsu, to discover, if possible, conclusive evidence in support of either Wrinch's or Bergmann's theory. The problem was developed in two ways. Dr. E. J. Wilson, Jr. prepared and studied long-chain peptides and obtained results which appear to support Bergmann's theory.

Dr. Trister stated that he had carried out the partial degradation of proteins to obtain and study large peptides and that his results to date also favor Bergmann's theory. In his experiments gelatin was added to a saturated solution of alcoholic hydrogen chloride and the mixture heated on a boiling water bath while hydrogen chloride was bubbled through until all of the gelatin went into solution (after approximately ten minutes). This solution was filtered, concentrated in vacuo, mixed with absolute methanol and precipi-

tated by being poured into absolute ether. The resulting white amorphous precipitate was removed by filtration and the hydrogen chloride was removed by dissolving the hydrolysate in methanol and adding an excess of silver oxide. After 24 hours' shaking a sticky mass of silver oxide and silver chloride had separated sufficiently from the supernatant liquid to allow the latter to be decanted. This was concentrated by evaporation, and precipitated by ether treatment. The amorphous substance obtained accounted for 35 per cent. of the starting material and was found to contain traces of silver.

The silver oxide-silver chloride residue was shaken with water. The resulting supernatant liquid was decanted, concentrated and treated with methanol, at which point a buff-colored precipitate appeared and was removed by filtration. This fraction accounted for about 50 per cent. of the starting material, and on analysis was shown to contain 5 per cent. silver. The other 15 per cent. of the original material was supposedly both alcohol- and water-insoluble and so remained in the silver oxide-silver chloride residue.

As yet, the nature of the linkage of the silver in the peptide has not been determined. It may be attached to an NH_2 group, a COOH group, or at some other position on the molecule. It was removed from the various fractions by hydrogen sulfide. By analysis, the total nitrogen of the alcohol-soluble fractions was found to be 14.34 to 14.38 per cent. in all instances, while the alcohol-insoluble fraction contained 15.3 to 15.4 per cent. nitrogen. From these analyses it was concluded that the difference in the fractions is real, and the treatment with silver oxide was responsible for the differentiation of the two products. Further experimentation may reveal information concerning the homogeneity of the various fractions, of which little is known at present.

Some preliminary experiments on the quantitative isolation of the individual amino acids of the alcohol-soluble peptide indicate that the same periodicity occurs in the alcohol-soluble peptide as in the gelatin molecule. Dr. Trister stressed the fact that a small error in the quantitative analysis of a protein would affect the whole number periodicity to a large extent. He cited specifically the work of Meyer who, on reinvestigat-

ing the amount of tyrosine in silk fibroin in 1940, found 10.8 per cent. instead of 13.2 per cent. as reported by Bergmann. This difference would change the periodicity of tyrosine from every sixth to every twentieth position in the silk fibroin molecule.

Following his talk, Dr. Trister discussed and answered questions asked by members of the group. He closed by expressing his thanks to Dr. Pacsu, with whom he has been working, for permission to include in his talk many unpublished results.

(Reported by MARY THOMAS ALLEN)

BOOK REVIEWS.

DIFFUSION IN AND THROUGH SOLIDS, by Richard M. Barrer, 464 pages, illustrations, 15 × 23 cms. New York, The Macmillan Company, 1941. Price $6.50.

The subject of this book is one of importance in many diverse lines of work and one in which research may yet solve important problems. The work is laid out with this in mind. In fact it is not only an exposition of data and information available but the effect is to point out and encourage possible fruitful lines of further investigation.

At the very outset a very definite need is met, that of giving some solutions of the diffusion equation in a form in which they may be applied, together with cases of diffusion systems in which the boundary conditions are those of the diffusion equation solved. This eliminates the special and often laborous technique in the treatment of equations of diffusion kinetics and enables the solutions to be applied at once. Stationary and non-stationary states of molecular flow in capillary systems is the next topic. As a background for this the different types of capillary flow are discussed such as molecular effusion, molecular streaming or Knudsen flow, Poiseville or stream-line flow, turbulent flow, and orifice flow.

With regard to gas flow in and through crystals and glasses the author recognizes that while the structure of alkali halides and similar more simple crystals is well known, it is not so with the large families of derivatives of silica. There are given a few of the known structural relationships for substances such as silica, silicate glasses zeolites, mica, clays, feldspars and ultramarines.

There is much information concerning the uptake and evolution of gases by metals. The treatment here covers the solubility of gases in metals and alloys, permeation velocities, the influence of temperature and pressure on permeability etc.—the main phenomena of gas flow through metals. This is supplemented with the study of diffusion of gases and non-metals in metals. Following this there is a description of the phenomena of conductivity and diffusion of ions and atoms in ionic lattices and metals. Both the reversible diffusion phenomena and the irreversible diffusion processes depending upon the past history of the system concerned are considered. Many of the problems discussed in the latter have to do with diffusion processes down grain boundaries, which lead to the treatment covered in the next chapter under the heading of migration in the surface layer of solids. The last part of the work deals with the permeation of gases and vapors with regard to organic solids.

The work is the result of an exhaustive search through the literature for authoritative references, the digesting of these references, coupled with expert knowledge of the subject and a keen sense of the art of presentation. It most certainly provides a means of enlightenment which would be next to impossible to obtain by a mere perusal of published data from a wide variety of sources. It is a stepping stone in helpfulness toward further research.

R. H. OPPERMANN.

COPYING TECHNIQUE, by Frank R. Fraprie and Robert H. Morris, 128 pages, illustrations, 16 × 24 cms. Boston, American Photographic Publishing Co., 1940. Price $1.50.

Documentary reproduction processes such as through the use of the photostat blueprint Recordak and other types depend for their success upon the technical skill of the worker and the quality of the orginal work. These are commercial processes. On the other hand, the copying of flat pieces such as engravings, paintings, book illustrations, etc., by means of photography brings into play individual treatment based on knowledge and initiative. The successful reproduction depends largely on the worker, the method he chooses and his personal application or modification of the method to his purpose.

The little book at hand discusses photographic copying methods by the use of the camera or a contact printing process. It first considers the equipment needed for copying in general and the preparations to work. The important questions of plates and films, exposure and development are taken up with regard to line work such as the pen and ink sketch and full tone pictures—both black and white. Copying colored originals such as oil paintings, water colors, color prints etc. is next considered followed by a consideration of the use of the infrared, ultraviolet and X-ray. Some space is given to specialized copying such as daguerreotypes, unfixed proofs, ferrotypes and ambrotypes, and expert photographic examination of documents. The last chapter is devoted to microfilm copying.

The compilers of the book recognize and so state that no book can give specific instructions as to the best way of obtaining copies. However, with a good understanding of what is given in the pages of this book the reader should profit.

R. H. OPPERMANN.

THE HARVARD BOOKS ON ASTRONOMY, edited by Harlow Shapley and Bart J. Bok; The Milky Way, by Bart J. Bok and Priscilla F. Bok; Earth, Moon and Planets, by Fred L. Whipple; Between the Planets, by Fletcher G. Watson; The Story of Variable Stars, by Leon Campbell and Luïgi Jacchia. About 250 pages to each volume, illustrations, 16 × 22 cms. Philadelphia, The Blakiston Company, 1941. Price $2.50 each.

Almost half of this fine series by members and associates of the staff of the Harvard College Observatory has now appeared. The showering of these four beautifully printed and elegantly bound books within the past two months is a boon to the average educated adult or adolescent reader who really is interested in modern astronomy. Under the editorship of Harlow Shapley and Bart J. Bok, we shall soon have completed a work on astronomy which will make it very difficult to recommend any other except for the very penurious.

Each of these books is exhaustive in its own field. The two books by Whipple and Watson respectively will answer almost any conceivable question about the solar system, even to where Mars will be, among the stars, on Franklin's Birthday in 1968! In each volume, the author has gone far out of his way to provide us with a wealth of unfamiliar and important material not to be found elsewhere except in the technical literature. Yet the presentation is always directed toward the non-technical reader, without insult to the vocabulary and intelligence of a professional. As a matter of fact, I believe one of the greatest uses of these books will be in the education of the professional in cognate fields. Too many

physicists, geologists, and chemists at some time or other dabble in narrow phases of astronomy without being fully informed in the astronomical knowledge and requirements of the situation. All physical and mathematical scientists should read several of these books, and let the popular books, written by hearsay by third-hand amateurs, strictly alone.

The illustrations, both half-tones and line drawings, are multitudinous and beautiful. How the authors (especially Whipple) extracted them from the original observers is baffling to one who has tried to get new pictures to freshen up an old story. Full tables where necessary, and good complete indices, distinguish all the books.

While Whipple and Watson were given the most difficult assignments, in view of the great wealth of material to be covered, there is no conspicuous difference in the homogeneity of their books, as compared with the other two, written by workers who spend most or all of their time in those very fields. The writing is in all instances more than sufficient, and in some passages, particularly in Bok-Bok, is beautiful.

To list detailed chapter headings would be of no value, for these HARVARD BOOKS. This reviewer can only assure potential readers that if they want to know anything about any phase of our modern astronomical knowledge covered by the titles of the books already at hand, they will find as much as is good for them to know. There remain five titles in the series, yet to appear. I hope they arrive soon, for I have read these first four from cover to cover with great enthusiasm, and I expect to do the same with those that follow. I'm sure I won't be disappointed, when I do.

R. K. MARSHALL.

BETWEEN PHYSICS AND PHILOSOPHY, by Philipp Frank, 238 pages, 14 × 22 cms. Cambridge, University Press. Price $2.75.

Philipp Frank, Visiting Lecturer, Harvard University, the author of this book was a member of a group of scientists known as the Vienna Circle who started about 1910 to reconstruct the ideas of Ernst Mach, the Austrian physicist, toward a movement for a clearer conception of the philosophy of science. This book contains nine papers in which Mr. Frank has from time to time expounded his thinking on the subject. He states that the principal aim of the papers is to show that use can be made of the recent progress of the physical sciences for confirming the stand that in contrast to the nineteenth century movement in line with materialism and naturalism, the twentieth century brings a trend away from these movements—an important sign of a general change in intellectual life. This is true only if interpretation is made of the progress according to the pattern of some cherished philosophy, disregarding the scientific meaning of physics.

A further aim is to make a contribution to the history of the development of the movement now generally called "logical empiricism" or "logical positivism." The first paper on the law of causality and experience, published in 1907, contains an idea of the opinions of an earlier group which the Vienna Circle superseded as regards positivism. How the teachings of Mach developed further is seen in the next paper on the importance of Ernst Mach's philosophy of science for our times, which appeared in 1917. The third paper devoted to physical theories of the twentieth century and school philosophy contains an exhibition of the synthesis

of positivism and presents explicitly the new logic. The connection of the European positivistic ideas with American pragmatism is clearly revealed. This paper was published in 1930.

"Is There a Trend Today Toward Idealism in Physics?" is the question discussed as the fourth paper. It was delivered at a conference in 1934 by the author and the following year he published paper number five on the standpoint of the positivistic conception of physics as opposed to the metaphysical conception represented in particular by Max Planck. At the First International Congress for the Unity of Science a series of papers were read. Among them was the sixth paper about the philosophy of the Soviet Union. Following this are papers on the philosophical misinterpretations of the quantum theory read by the author in 1936 at the second Congress for the Unity of Science, a discussion of what "length" means to the physicist (1937), determinism and indeterminism in modern physics (1938), and Ernst Mach and the unity of science (1938).

The book very definitely establishes a trend of thought, an evolution, which not only is interesting but forms a basis for speculation and a means of testing and application as to its soundness as science progresses.

R. H. OPPERMANN.

PUBLICATIONS RECEIVED.

Standard Handbook for Electrical Engineers, by Archer E. Knowlton, Editor-in-Chief, 2303 pages, illustrations, 16 × 23 cms., New York, McGraw-Hill Book Co., Inc., 1941. Price $8.00.

The Microscope, by Simon Henry Gage, 617 pages, illustrations, 16 × 24 cms., New York, Comstock Publishing Company, Inc., 1941. Price $4.00.

Diffusion in and through Solids, by Richard M. Barrer, 464 pages, illustrations, 15 × 23 cms., New York, The Macmillan Company, 1941. Price $6.50.

Science Calls to Youth, by Raymond F. Yates, 205 pages, illustrations, 14 × 20 cms., New York, D. Appleton-Century Company, 1941. Price $2.00.

United States Department of the Interior, Bureau of Mines. Minerals Yearbook Review of 1940. 1459 pages, 16 × 24 cms. Washington, D. C., Government Printing Office, 1941. Price $2.00.

The Harvard Books on Astronomy, edited by Harlow Shapley and Bart J. Bok; The Milky Way, by Bart J. Bok and Priscilla F. Bok; Earth, Moon and Planets, by Fred L. Whipple; Between the Planets, by Fletcher G. Watson; The Story of Variable Stars, by Leon Campbell and Luigi Jaccia. About 250 pages to each volume, illustrations, 16 × 22 cms. Philadelphia, The Blakiston Company, 1941. Price $2.50 each.

Industrial and Engineering Chemical Technology, by Francis Wm. Brown, Ph.D., 102 pages, 22 × 28 cms. Philadelphia, Stephson-Brothers, 1941. Price $1.50.

Temperature Measurement and Control, by Robert L. Weber, 430 pages, illustrations, 15 × 22 cms. Philadelphia, The Blakiston Company, 1941. Price $4.00.

Fatigue of Workers, Its Relation to Industrial Production, by Committee on Work in Industry of the National Research Council, 165 pages, 16 × 24 cms. New York, Reinhold Publishing Corporation, 1941. Price $2.50.